CLAREMONT GRADUATE UNIVERSITY

Peter F. Drucker and Masatoshi Ito Graduate School of Management

AN EMPIRICAL STUDY OF A NO-ARBITRAGE LIQUIDITY MODEL IN FINANCIAL MARKETS WHERE LIMIT ORDER BOOKS ARE MODELED BY A BROWNIAN SHEET

A Dissertation in

Ph.D. in Management Program

by

THANH HOANG

A Dissertation submitted to the Faculty of Claremont Graduate University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the Graduate Faculty of Management

(*with a concentration on Financial Engineering*)

Approval of The Review Committee

This Dissertation has been thoroughly read, reviewed, and critiqued by the Committee listed below, which hereby approves the manuscript of THANH HOANG as fulfilling the scope and quality requirements for meriting the degree of Doctor of Philosophy in the Graduate Faculty of Management with a concentration on Financial Engineering.

HENRY SCHELLHORN, Chair

Director of the Institute of Mathematical Sciences

Professor of Mathematics

BERNIE JAWORSKI

Interim Dean, Peter F. Drucker Chair in Management and the Liberal Arts

Professor of Management

QIDI PENG

The Institute of Mathematical Sciences

Professor of Mathematics

Abstract

An Empirical Study of a No-Arbitrage Liquidity Model in Financial Markets where Limit Order Books are modeled by a Brownian Sheet

by

Thanh Hoang

Claremont Graduate University: 2013

We present an empirical study of a no-arbitrage liquidity model in financial markets.[1] Our research approach differs from previous liquidity studies which focus on a price-driven market to analyze solely the dynamics and characterization of the limit price process. In our research, we concentrate on an order-driven market where the equilibrium prices (also called *clearing prices*) of the assets are completely determined by the limit order flow. By employing a Brownian sheet and the Ito-Wentzell formula, we model the net demand curve and the clearing price process, as well as characterize their dynamics under an equivalent measure. We discover that the volatility of the clearing price process is inversely

[1] The no-arbitrage dynamic liquidity model for the limit order book was developed by Professor Henry Schellhorn and Professor David German with the collaboration of Thanh Hoang. The paper has been submitted and reviewed at SIAM Journal of Financial Mathematics in June 2012.

proportional to the sum of buy and sell limit order flow density. We develop a "semi-relative" no-arbitrage model of liquidity[2] for our empirical analysis. By using high-frequency aggregate limit order real-time data from NYSE Arcabook[3] for four different industries with high-frequency trading, including energy, financial banking, materials and mining, and technology; we simulate our model to approximate the price and calculate the implied volatility for each European option. Our simulations show that in our model of liquidity, a volatility smile arises spontaneously. Since volatility smiles are observed in real markets, this confirms the plausibility of our model.

[2] Our model is different from the no-arbitrage model of liquidity with relative prices, presented in German and Schellhorn (2012).

[3] Historical NYSE Arcabook real-time data are recorded the complete information of limit order transactions entered into the ArcaEx or ArcaEdge system from OTC, exchange-listed, ETF and OTCBB securities.

Acknowledgments

I am heartily thankful and truly indebted to a number of people who had provided me with their endless support, guidance, and commitment. There is no doubt that my Ph.D. degree in Management with a concentration on Financial Engineering is one of the most challenging missions which I have attained in my life. During the past five years in the Ph.D. program at Claremont Graduate University (CGU), one of my greatest pleasures here had been the associations and opportunities I have to meet and be friend with varieties of people from all walks of life, thus help enrich my experience and I learned a lot from each one of them. I will cherish all the help and support I have received from this long but worthwhile road in my educational advancing endeavor.

First and foremost, I would like to thank all of my professors at CGU. Most distinguishably, Professor Henry Schellhorn, my mentor and chair of advisor, whom I am forever indebted to. Professor Schellhorn plays a vital roll in my education journey at CGU. It has been an honor and pleasure for me to have worked with him during these past years. His insightful advices and advance visionary is not only admirable but significantly impact this research by maneuvering it in both theoretical and practical aspects. The commitment and rigorous work for this research requires diligent, comprehensive understanding, and advance skills in both mathematic and programming to make it transpire. He is able to guide me through a havoc web of data and a complicated model with his encouragement, unlimited support, and provided me the freedom to explore independently.

I would like to express my sincere thanks to my committee, Professor Bernie Jaworski and Professor Qidi Peng for their help and support. I would like to thank Professor David German, who provided the materials, support, and encouragement. I am also grateful to Professor John Angus, who accepted and guided me when I started my Ph.D. studies at CGU. To Professor Andrew Nguyen and Professor Vijay Sathe for their mathematical and social science advices. I also greatly appreciate Professor Murat Binay, Professor Hedley Morris, Professor Robert Pedace, Professor Alpan Raval, Professor Richard Smith and other professors, who have faith and encourage me throughout my studying at CGU.

Most importantly, I would like to express my heartfelt gratitude to my parents for their supports emotionally and financially. And to recognize a very special person who has been by my side all these years, my wife Vanessa. Her unconditional love gave me faith to face some of the most difficult obstacles in my life, while her generosity of understanding, patience, and encouragement gave me strength to overcome my degree challenges.

Finally, as a member of Claremont Graduate University, I have been surrounded by wonderful people whom I will never forget. I would not have achieved my dream without the encouragement from all of my friends, and the time my professors put in to help enrich my learning experience. I must thank Adrian Escalera and Anthony Gonzalez for their assistance, both had been supportive and helped me for the research data. I must also thank those who may not have contributed directly but emotionally wished me well; my friends, Jay Hong Kim, James Preciado, Albert Ramos, Sudha Sethuraman, and Tanakorn Termpongnurak, thank you very much for your encouragement. To my friends at the stochastic group, thank you. It had

been an honor and my privileged to have received so much love that makes my degree a reality, for that, the debt of my gratitude to you all is immense.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Over the last decade, the evolution of high-frequency trading (HFT) has taken the financial world by storm. Taking advantage of the relentlessly fast-paced markets and supercomputers, high-frequency (HF) traders implement highly quantitative computerized algorithms to analyze the incoming market information in a matter of microseconds and employ sophisticated trading algorithms to manage liquidity risks and thus make profits. Moreover, the regulations of Securities and Exchange Commission (SEC) had changed in recent years, which also contributed to an enormous popularity of HFT. For example, in 2001, the change of quoting prices in decimals instead of fractions[1] resulted in traders finding better alternative trading methods to make profits, which in turn gave a strong support leading to increase the important role of HFT. In 2005, the Regulation National Market System (NMS) passed by the SEC sustained the transparency, consistency and competition between markets by promoting fair and non-discriminatory access to NMS trading centers and providing some new reinforcement and protection rules. Therefore, under the new regulation, traders could profit from

[1] The U.S. stock exchanges began bringing down the minimum spread between the buy and sell prices from 1/6th of a dollar to one cent in 2001. *Source*: SEC, Bloomberg.

any small difference in price of any security between two different stock exchanges by taking advantage of the time lag.[2]

Accounting for approximately 51% of daily equity trading volume in the United States stock markets in 2012 compared to 21% in 2005, HFT is playing a crucial role in the financial trading system with a compounded annual growth rate of 22% in trading volume.[3]

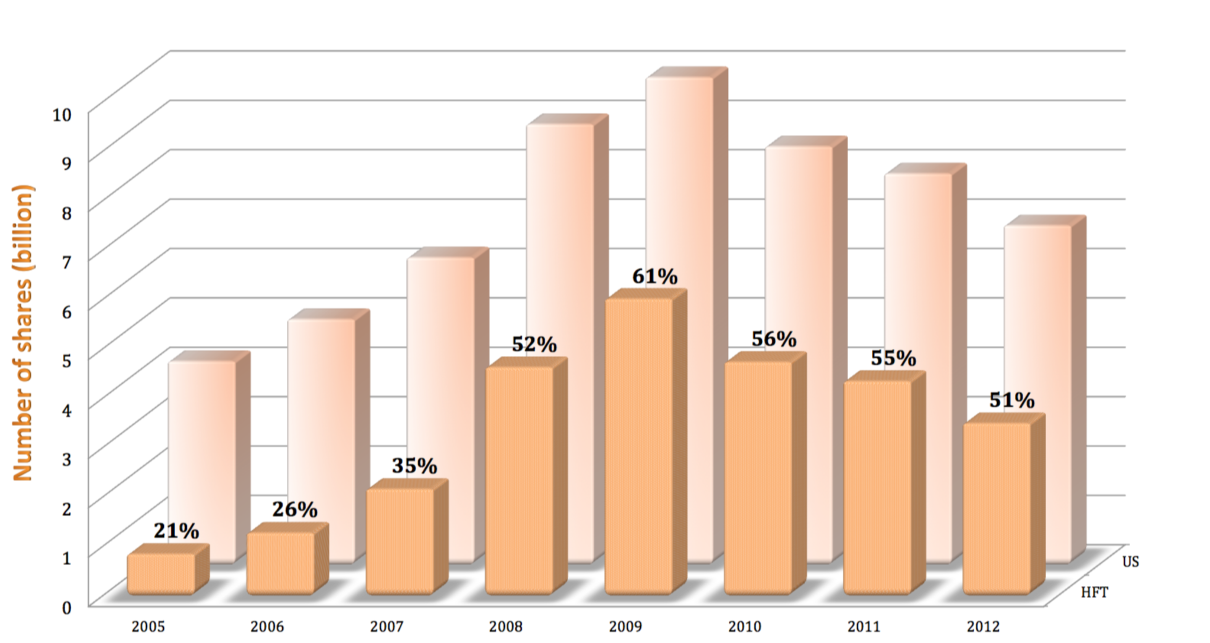


Figure 1.1: Average daily trading volume by HFT firms in all U.S. stocks (2005-2012).

Although HFT may increase the price fluctuations and short-term volatility,[4] there is no doubt that the large trading volume entered by HF traders helps creating greater liquidity in the market, which in turn not only ameliorates the liquidity level, but also improves market efficiency because the reflected prices are now more accurate due to the instant updated information from the stock exchanges. With that said, the impact of HFT on the U.S. market is still considered as a dilemma, which needs more time to be investigated and is out of the

[2] SEC Release No. 34-51808; File No. S7-10-04, *effective date*: August 29, 2005.

[3] The data are collected from Agarwal (2012), combined with other sources including Tabb Group, Rosenblatt Securities and The New York Times.

[4] For example, some financial experts believe that the "flash crash" on May 6, 2010 was contributed by the behavior of HF traders.

scope of our research.[5]

Regarding the liquidity on the stock exchange, traders (e.g. market makers or financial firms) control it by submitting their orders, including buy and sell orders. In general, there are two main types of orders, such as (1) market orders where these orders are executed at the best available price (*immediate price*) and (2) limit orders in which the trader sets a price which he/she is willing to trade if a counter-party meets that price (*limit price*). Benefiting from an open access to current real-time limit order books (LOBs) of stock exchanges (e.g. NYSE, NASDAQ), most HF traders implement models on the bid-ask spread of limit orders to make profits. In this strategy, regular HF traders generally place their limit orders to buy slightly below the current market price or sell slightly above the current market price to ensure that they will achieve their better prices.[6] Conversely, large HF traders usually place their limit orders to buy at higher current market prices or sell at lower current market prices due to the liquidity capacity of the assets. Consequently, the stock exchange system can automatically execute these limit orders according to precise time and limit price priority rules. Therefore, the challenge of understanding the dynamics of the LOBs with HF intraday data attracts a lot of interests in mathematical modeling, motivated by not only HF traders, but also academic researchers.

1.1 Limit Order Matching Mechanism in LOB

In mathematical finance, many liquidity models concentrate their studies solely on the dynamics and characterizations of limit price process, such as Parlour (1998), Sandås (2001), Avellaneda et al. (2011), and Zheng et al. (2012). In other words, their research studies focus on a price-driven market. Our research, however, approaches a model of liquidity in an

[5] The chairman of the SEC Mary L. Schapiro confirms that until now, we do not have enough information (i.e. data) to understand in-depth the impact of HFT on the U.S. markets. *Source*: Hilzenrath (2013).

[6] *Source*: http://www.investopedia.com.

entirely different way. In our model, we consider that the equilibrium prices, defined as the clearing prices $\pi(t)$ of the assets, are completely determined by the limit order flow. We model a market of assets in the LOB without a specialist, where buyers and sellers submit their limit orders at HF. In other words, our research focuses on an order-driven market. For a buy limit order, the trader places an order of a given quantity of an asset at a given buy limit price defined as the *maximum price* at which he/she is willing to pay. For a sell limit order, trader places an order of a given quantity of an asset at a given sell limit price defined as the *minimum price* at which he/she is willing to sell.

(a) Unexecuted limit orders in LOB at $0 < t < 1$,

(b) A new buy limit order arrives at $t = 1$,

(c) Limit order matching in LOB at $t = 1$,

(d) LOB at $t = 1$ with $\pi(1) = \$34.20$.

Figure 1.2: Limit order matching mechanism in LOB.

For example, at time $t > 0$, the LOB includes unexecuted buy and sell orders arranged by their time of arrival, which are awaiting execution with different price levels. At time $t + \Delta t$ where $\Delta t > 0$, a trader submits a new buy limit order into the exchange system. His/her buy limit order is matched with an existing sell limit order at the required limit price, but he/she is unable to complete his/her entire order due to the shortage of outstanding sell limit orders. Therefore, the unexecuted remaining part of his/her order will be recorded in the LOB.

By following the limit order matching mechanism in Figure 1.2, new incoming buy limit orders may be matched with the outstanding sell limit orders, that previously existed in the LOB. A symmetric outcome also follows in case of new coming sell limit orders. In the HFT system, time is of the essence; thus the time when the order is placed, is a tie-breaker among new incoming buy limit orders as well as among new incoming sell limit orders. As a result, under this matching mechanism, the clearing price process $\pi(t)$ is always defined and determined solely by the limit order flow.

Example 1.1.1. We suppose that the clearing price of Oracle Corporation at time $t = 0$ can be any price in the range from \$34.01 to \$34.43 on April 4, 2011[7], i.e. $\pi(0) \in [34.01, 34.43]$. During $0 < t < 1$, the LOB contains the following unexecuted limit orders.

Buy Price	**Buy Quantity**
\$34.05	50
\$34.10	100
\$34.15	120

Sell Price	**Sell Quantity**
\$34.20	50
\$34.30	100
\$34.35	150
\$34.40	80

Table 1.1: Unexecuted limit orders in LOB, $0 < t < 1$.

At time $t = 1$, a trader submits a buy limit order of 150 number of shares at a buy limit price of \$34.25. The stock exchange matches this order with existing unexecuted sell limit orders, i.e. the sell limit order of 50 number of shares at a sell limit price of \$34.20. However, his/her

[7] We assume the price range based on the highest and lowest historical prices of Oracle Corporation (ORCL) on April 4, 2011. *Source*: Yahoo! Finance.

order can only be filled partially, and the remaining part of his/her order will be recorded in the LOB.

Buy Price	Buy Quantity
\$34.05	50
\$34.10	100
\$34.15	120
\$34.25	100

Sell Price	Sell Quantity
\$34.30	100
\$34.35	150
\$34.40	80

Table 1.2: Unexecuted limit orders in LOB, $t = 1$.

Consequently, the clearing price at time $t = 1$ is equal to the limit price of the sell order, i.e. $\pi(1)$ = \$34.20. This example illustrates some properties of the limit order markets:

1. Under this matching mechanism, the clearing prices $\pi(t)$ is always defined, which we can assume any positive value.[8]

2. It is common to think how an incoming order "crosses" the LOB. Crossing the LOB has two advantages:

    - First, its execution is faster. For example, if the trader submitted a buy limit order with its price of \$34.30, his/her buy order with a quantity of 150 number of shares would be completed as he/she desired, rather than waiting an indeterminate amount of time until suitable sell limit orders arrive at his/her buy limit price.
    - Second, we suppose that traders submitted some buy limit orders at the same time when the demand exceeds the supply at the lowest sell order limit price, the buy limit order with the highest limit price is executed first. This is consistent with the theory of optimal LOB placement presented by Roşu (2009).

By following the above matching mechanism, we distinguish two different types of limit orders in the trading system, including *immediate orders* and *wait orders*. Orders of the former

[8] We do not consider markets for swaps, where the prices can be negative.

type are placed at the limit prices, so that they are more likely to be instantaneously matched with existing limit orders according to precise time and limit price priority rules. On the other hand, orders of the latter type need to spend a positive amount of time in the LOB before they are matched, modified or canceled.

1.2 The Absence of Market Manipulation Strategies

In our model, we assume that the absence of market manipulation strategies in discrete-time holds under Jarrow (1994)'s condition, and our approach takes advantage of the results of German and Schellhorn (2012) which proved the absence of arbitrage in continuous time under certain conditions. We also consider that the LOB data are public information. This assumption is both theoretically and empirically reasonable because in practice, the automated HFT has largely replaced floor-based trading in equity markets from the above evidences in Figure 1.1. Furthermore, with the rapid development and expansion of Electronic Crossing Networks, established stock exchanges such as NYSE, NYSE Euronext, NASDAQ, the London Stock Exchange, and the Tokyo Stock Exchange etc. have fully or partially provided electronic order-driven trading systems.[9] As a result, HF traders can easily access in the real-time LOB data from these stock exchanges, considered as public information. We will continue to discuss the market manipulation strategies more in-depth in Chapter 2.

The structure of this Dissertation is organized as follows. In Chapter 1, we introduce briefly the HFT system, limit order matching mechanism, limit order types and some basic assumptions from our model. In Chapter 2, we present a literature review of liquidity under three main categories: (1) liquidity in general, (2) market liquidity and (3) liquidity models. In Chapter 3, we define our model in details, including our assumptions, definitions, lemmas

[9] Electronic Crossing Networks including Archipelago, Instinet, Brut and Tradebook etc. provide HF traders the complete information of order-driven trading orders. *Source*: Cont (2011).

and theorems. In Chapter 4, we discuss our semi-relative no-arbitrage model of liquidity and the implementation of the model for the empirical analysis, present the structure of NYSE Arcabook data[10], and our data handling process and security selection with at least two different stocks from four different industries[11] for the empirical analysis. In this chapter, we also present the simulation methodology to approximate the price and calculate the implied volatility for each European option under certain conditions. In Chapter 5, we show our simulation results in two groups: (1) favorable results, and (2) open issues for extensive studies. In Chapter 6, we summarize our theory, make our conclusions based on our simulation results and suggest some extensions to our research.

[10] NYSE Arcabook provides the complete LOB data from NYSE, NYSE Arca, NYSE MKT, NASDAQ and the ArcaEdge platforms. The data are collected every trading day from 3:30 a.m. to 8:00 p.m. ET. There are more information at http://www.nyxdata.com.

[11] We not only select HFT stocks in very active and volatile industries, such as energy (CVX, XOM), materials and mining (ABX, FCX), but also examine some stocks in the financial banking (JPM, WFC) and technology industries (CSCO, MSFT, ORCL).

Chapter 2

Literature Review

In this chapter, we summarize the literature of liquidity studies in three main categories: (1) liquidity in general, (2) market liquidity and (3) liquidity models.

2.1 Liquidity

The concept of liquidity in finance was popularized by Keynes (1936) in his General Theory. Keynes presents that the liquidity of money trading, including money demand and supply, determines the interest rate. He believes the aggregate demand of the economy does not necessarily equal its aggregate supply. In other words, the difference between the aggregate demand and supply creates the liquidity capacity of an asset. He also defines the liquidity as "more certainly realizable at short notice without loss". Hicks (1962) explains this definition that the liquidity is defined as a marketable asset's ability to be sold quickly without the loss of its value, and not depended on the up-trends or down-trends of the market. Hirshleifer (1972) specifies the liquidity capacity of an asset as the fund available for immediate consumption or reinvestment in the form of money. These concepts are the antecedents of the current liquidity terminology in finance.

In the equity trading field, liquidity contributes a significant impact on the stock prices and their rates of return, but liquidity is not easy to be observed directly. Therefore, we generally summarize two important major measurements of liquidity including (1) the quoted bid-ask spread, (2) trade sizes and transaction costs.

In the first category of the quoted bid-ask spread as a measure of liquidity, Amihud and Mendelson (1986) prove the positive return-illiquidity relationship. They also suggest an increase in the firm's liquidity financial policies may reduce the opportunity cost of capital, which in turn enhances the value of the firm. Easley et al. (1996) sustain this belief because it is obvious that illiquid stocks are riskier than liquid stocks, and logically yield higher rates of return. Eleswarapu and Reinganum (1993), however, point out the impact of the relative bid-ask spreads on stock returns from Amihud and Mendelson (1986)'s study is only significant in the month of January during the 1961-1980 period while it has no effect in non-January months. However, both of these studies implement the closing bid-ask spreads as a proxy for the liquidity measurement without consideration of the length of the holding period over which these spreads are amortized. As a result, Chalmers and Kadlec (1998) propose the amortized spread as a direct measure of a stock's liquidity. The amortized spread of a stock is equal to the product of its effective spread and share turnover in order to measure the annualized cost of the spread to investors. They discover that the amortized spread of a stock has a negative correlation with the firm's market values of equity while it has a positive correlation with its book-to-market. In general, stock returns are negatively correlated to market value of equity[1], and positively correlated to book-to-market[2]. However, it is a surprise that they detect a weak support for a cross-sectional relation between stock returns and amor-

[1] Banz (1981) shows that during the 1936-1975 period, on average, the risk-adjusted stock returns of small firms with small total market value of equity are higher than those of large firms with large total market value of equity. Reinganum (1981) supports this opinion to find with equivalent risks, the average rates of return of small firms are higher than those of large firms during the 1967-1975 period.

[2] Rosenberg et al. (1998) supports this point of view by examining a sample of 1,400 largest companies in the Computstat database during the 1980-1984 period.

tized spreads. Alternatively, they notice amortized spreads have a positive correlation with the volatility of stock returns. With the limitations and inconsistency of the bid-ask spread as a measure of liquidity[3], other researchers believe that liquidity could be captured with a different method by using the variable components of trading sizes and transaction costs.

In the second strain of research, by using the transaction data, particularly trade sizes and transaction costs as a measure of liquidity, Brennan and Subrahmanyam (1996) conclude that illiquid stocks have significant positive correlation with their monthly returns. In another way, Breen et al. (2002) measure the equity liquidity by estimating the change in the company's stock price with its observed net trading volume, while Amihud (2002) proposes the illiquidity measure as the average daily ratio of absolute stock return to dollar volume in order to detect the positive effect of illiquidity on stock returns, consistent with earlier studies. O'Hara (2003) figures out the relationship between liquidity and price by referring liquidity as the matching between buyers and sellers in the market, which in turn creates the asset price formation and then impacts on the price discovery process. In addition, by applying the liquidity adjustment, Acharya and Pedersen (2005) conclude the liquidity-adjusted CAPM[4] is better than the standard CAPM. Their model also shows that liquidity predicts future returns and its move in a correlation with returns, which is consistent with Amihud (2002).

For the current definitions of liquidity, Brunnermeier and Pedersen (2009) define liquidity in finance in two different forms, i.e. (1) trading liquidity (how easy an asset can be traded in markets), and (2) funding liquidity (how easy a trader can obtain funding). In our research, we focus on the former type of liquidity, the trading liquidity. Overall, liquidity in trading fi-

[3] Grossman and Miller (1988) point out some limitations of the bid-ask spread as a measure of liquidity: (1) the bid-ask spread may present a special case where market makers act for both sides of buyers and sellers, (2) the bid-ask spread does not contain the proportion of risk when buy and sell orders are separated in time, (3) transaction costs are highly related to the trading risk when traders may have no way of recovering these costs to make profits.

[4] Capital Asset Pricing Model.

nance describes as an asset's ability to be traded quickly at any time in the markets with a minimal loss of value. On the other hand, illiquid assets are difficult to trade because of their expense, limited quantity or lack of demand in the economy.

2.2 Market Liquidity

In the market liquidity literature, Black (1971) defines a liquid market as a continuous market where almost any amount of equity can be traded immediately at a very closed market price, i.e. small bid-ask spreads, while large amounts of equity can also be traded over a long period of time at a price, on average, not much different from the current market price. For example, if a trader executes a large buy order, it will reasonably affect the market price for a significant increase. Vice versa, if a trader executes a large sell order, the market price will decrease. However, over a long period of time, it is possible for equal chances of buying and selling; thus on average, the execution of large trades may not change the price as much. By holding Black's conditions, Kyle (1985) presents three characteristics to judge the market quality, including (1) market tightness measured by the size of bid-ask spreads; (2) market depth measured by the ratio of trading volumes over market-price change; and (3) market resilience measured by the speed at which the trade impact disperses. Kyle also shows that given the information available to the noise traders, the resulting price process is a martingale in the appropriate measure, whereas it may not be for the informed traders.

On the other hand, Grossman and Miller (1988) postulate that market liquidity is determined by the demand and supply of immediacy from market makers and outside customers.[5] In the long run, the equilibrium of the supply and demand for immediacy determines the level of market liquidity. Holmström and Tirole (1993) believe the market liquidity de-

[5] Outside customers will participate in the trading market when they observe a gap between their desired price and the current price, called as a liquidity event.

creases when the concentration of ownership increases, supported by the empirical results of Bhushan (1989). Holmström and Tirole (1998) also investigate the role of government in supplying and managing the market liquidity, compared to the private supply of liquidity. They discover in general, individual firms are not able to satisfy their liquidity needs by holding only private market instruments, which in turn creates a demand for the public supply of liquidity from the government (i.e. Treasury bonds). They point out that the important role of the government is not only in the supply of the market liquidity, but also in the active management of the market liquidity.

Alternatively, Chordia et al. (2001) measure the market liquidity by the quoted effective spreads with the market depth and the trading activity by the trading volume and the number of daily transactions to figure out how the market performance affects the trading on a given period of time. They conclude both the market liquidity and the trading activity are affected by equity market returns, recent market volatility, short-term interest rates and the term spreads. They also notice the market liquidity decreases and the trading activity slows on Friday while the opposite situation occurs on Tuesday. For the market liquidity of the U.S. bond markets, Chakravarty and Sarkar (1999) discover the realized bid-ask spread as a measure of liquidity decrease in the trading volume in all corporate, municipal and government bond markets during the 1995-1997 period. Huang and Wang (2009) investigate the impact of the stock market liquidity on asset prices to figure out the relationship between the market liquidity and market crashes. They notice the shortage of market liquidity always decreases the stock price significantly, and which in turn causes a market crash. They also point out that market illiquidity leads to high expected returns of stocks, consistent with earlier studies.

In our research, we do not intentionally analyze the market liquidity, but in fact, we implement the definition of a liquid market as a continuous market (see Black, 1971) to assume that the net demand curve Q will be continuous in time, and under certain conditions (see Kyle, 1985 and German and Schellhorn, 2012) to prove that the clearing price $\pi(t)$ is a martingale in the filtration corresponding to the public information, and not private information.

2.3 Liquidity Models

We generally classify the literature of liquidity models into two different categories: (1) large traders can manipulate the prices in the market, (2) all traders are considered as price-takers.

2.3.1 Large Traders as Market Manipulators

By definition, a large trader is any investor whose trades change market prices because of his/her trading quantity or because *with some probability*, other traders in the market believe that the large trader has been informed. Hart (1977) shows under the unstable stage of the dynamic economy, there always exist profitable opportunities for speculators; or even under the stable stage of the dynamic economy with certain conditions, manipulation is also possible. Jarrow (1992) extends this research from an infinite horizon, deterministic economy with a time homogenous price process to a generalized stochastic economy with a time dependent price process. He shows under reasonable hypotheses on the equilibrium price process expressed as a function of the manipulator's trading strategy, large traders with their market power can manipulate prices to make profits at no risk. On the other hand, he also provides some sufficient conditions for the non-existence of market manipulation.

Jarrow (1994) extends his earlier research to investigate the impact of market manipulation strategies on derivative security markets. In a discrete-time model, he shows that large traders usually employ two main strategies to manipulate market prices: (1) corner the market by obtaining sufficient amount of stocks, commodities or other assets in a particular industry without the illegal practice of monopoly; thus manipulate market prices to make profits; (2) front run their own trades by taking advantages of their advance information of pending orders from their customers to make profits.[6] In discrete time, Jarrow (1992) also shows that there is no market manipulation strategy if there is no-arbitrage for atomistic traders in periods when large traders do not trade. In our research, we employ the assumption of the absence of the market manipulation in discrete time as Jarrow (1994)'s conditions while we also prove and employ the absence of the market manipulation in continuous time under certain conditions.

In another way, Bank and Baum (2004) develop a general model of liquidity in the continuous time in an illiquidity financial market with a single large trader whose trades can change the market prices. They also employ the Ito-Wentzell formula to prove the absence of arbitrage for the large trader and conclude that the large trader can obtain the same utility as atomistic traders in the market. In our research, we take advantages of the results of Bank and Baum (2004) to implement in our continuous-time model for a market of atomistic traders and one large trader, which we will discuss in-depth in Chapter 3.

2.3.2 All Traders as Price-Takers

The second class of models considers all traders as price-takers under the absence of market manipulation, for example, Çetin et al. (2004), Çetin and Rogers (2007)[7], Çetin et al. (2010)[8],

[6] In general, the large number of submitted orders will predictably affect the price of the security. Therefore, front-run traders will buy or sell in advance for their own accounts before filling customers' orders to make profits at no risk. For example, a front-run trader either (1) buys for his own account before filling customers' buy orders, that will predictably drive the price of the security up, or (2) sells for his own account before filling customers' sell orders, that will predictably drive the price of the security down.

[7] Çetin and Rogers (2007) develop a model of liquidity to investigate the maximization of expected utility of terminal wealth. They prove that in an illiquid markets, optimal portfolios and arbitrage opportunities could co-exist and any liquidity costs (even small) could affect significantly on the hedging strategy.

[8] From the technical viewpoint, Çetin et al. (2010) sustain the belief that traders should make consecutive small trades instead of a large trade in an illiquid market. By applying the Ito's lemma, they invent a super-hedging strategy to overcome the trade-off between a higher-cost buy-and-hold strategy without liquidity costs and the Black-Scholes replicating strategy with liquidity costs by dividing the value of the option into two parts and hedging them in two different above strategies.

and Gökay and Soner (2011)[9].

Considering all traders who act as price-takers, Çetin et al. (2004) model a stochastic supply curve as a function of trading quantity. By including liquidity risk and studying an economy where the market price process depends on the trading quantity, they extend classical arbitrage pricing theory by showing that derivative prices are equal to the classical no-arbitrage price of the derivative in an approximately complete market. Since market orders are matched instantaneously in the trading system, Çetin et al. (2004) assume the residual supply curve at a future time is statistically independent from the order just matched. However; in our research, we consider this assumption as implausible because of two reasons: (1) all information is assumingly contained in the order flow; and (2) this assumption does not explain how prices can incorporate information from the arrival of new orders.

In the next chapter, we will present our model in two different scenarios: (1) a market with atomistic traders, (2) a market with atomistic traders and a large trader. We also define the clearing price process under the risk-neutral measure.

[9] Gökay and Soner (2011) develop a hedging strategy for European and barrier options in a binomial illiquidity market by proposing a sufficiently small liquidity parameter for hedging with a premium cost.

Chapter 3

The Model

Let us define a Brownian sheet $W(t,s)$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $\{\mathcal{F}_t\}$

$$W(t,s) = \sum_{j=1} \beta_j(t) \int\limits_{0 \leq \alpha \leq s} g_j(\alpha) d\alpha, \quad \text{for } t \geq 0, \text{and } s \in [0, S]$$

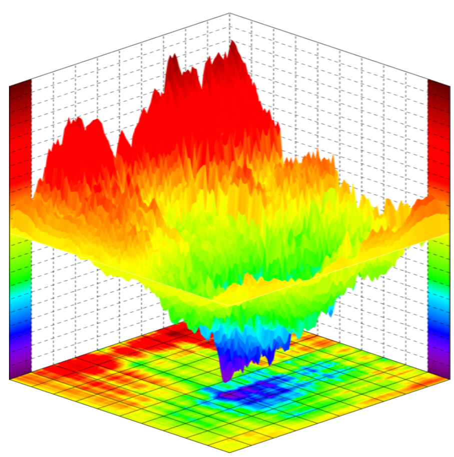


Figure 3.1: The dynamics of the Brownian sheet.
Source: Université de Poitiers - Mathématiques.

$\{\beta_j\}$**:** a list of independent real-valued standard Wiener processes,

$\{g_j\}$**:** an orthonormal and complete basis for some Hilber space,

$\{s\}$**:** the "factor" information, identified by the variable limit price p, where $p \in [0, S]$,

$\{\mathcal{F}_t\}$**:** the filtration, generated by the collection of Wiener processes $\{\beta_j\}$.

A stochastic integral of an $\mathcal{F}_t$-adapted integrand $\sigma(t,s)$ with respect to the Brownian sheet is defined by:

$$I(T,S) = \int_0^T \int_0^S \sigma(t,s) W(dt,ds).$$

Assumption 3.0.1. *Buy and sell limit prices can assume any real value between 0 and S. They are usually denoted by p. Orders can be submitted to the market at any time $t \in [0,T]$.*

We can assume the buy and sell limit prices to be any real positive value between 0 and S because we do not consider markets for swaps, where the prices can be negative. Under this assumption, the clearing price $\pi(t)$ determined by the order flow is also any real positive value between 0 and S, where:

$$\pi(t) = \begin{cases} 0 : \text{bankruptcy,} \\ S : \text{maximum limit price, set by the exchanges to prevent excessive speculation.} \end{cases}$$

Assumption 3.0.2. *The market is frictionless, i.e. $c = 0$.*

This assumption is standard in mathematical modeling, i.e. no transaction cost. By definition, the transaction cost is an addition fee, charged by the stock exchange for each transaction on top of the limit price. For example, a trader who submits a buy limit order of 50 number of shares must pay a fee for this transaction when his/her order is filled; thus eventually he/she has to pay $\pi(t)$ multiplied by 50 number of shares and plus c, where: $c > 0$.

3.1 A Market with Atomistic Traders

Definition 3.1.1. *The net demand curve Q is a function $[0,S] \times \mathbb{R}^+ \times \Omega \rightarrow \mathbb{R}$, which value $Q(p,t,\omega)$ is equal to the difference between the quantity of shares* ***available*** *for purchase and the quantity of shares* ***available*** *for sale at price p and at time t. For each p the stochastic process $Q(.,t,.)$ is $\mathcal{F}_t$-adapted. The random variables $Q(p,t,.)$ are assumed to be uniformly bounded in p and t almost surely.*

We assume that the net demand curve $Q(p, t)$ will be continuous in time, which represents the available information in the LOB of the stock exchange. In our model, we do not need to specify whether the process $Q(p, t)$ presents the net demand just before clearing or after clearing because these quantities move continuously in the HFT system.

Definition 3.1.2. *The clearing price $\pi(t)$ is an $\mathcal{F}_t$-adapted stochastic process which either satisfies*

$$Q(\pi(t), t) = 0, \tag{3.1}$$

when there is a solution to (3.1), or is otherwise defined by continuation, i.e., $\pi(t)$ is equal to the value of π at the latest time $s < t$ for which there was a solution to $Q(\pi(s), s) = 0$.

Since we assume the net demand curve $Q(p, t)$ is continuous in time, we have

$$Q(\pi(t), t) = Q(\pi(t), t_+), \quad \forall t.$$

Definition 3.1.3. *A limit order submitted at time t crosses the market if:*

- *either it is a buy order with limit price $p > \pi(t_-)$*
- *or it is a sell order with limit price $p < \pi(t_-)$*

We call these orders immediate orders. All the other orders are called wait orders.

From Definition 3.1.3, we can develop a market model in a general way, i.e. accumulate all buy order quantities with limit price higher than $\pi(t_-)$ which are submited into the exchange system and subtract the cancelled quantities when orders are cancelled to present a cumulative demand curve $\mathcal{D}(p, t)$; and similarly accumulate all sell order quantities with limit price lower than $\pi(t_-)$ which are submited into the exchange system and subtract the cancelled quantities to present a cumulative sell curve $\mathcal{S}(p, t)$. Therefore, we have

$$Q(p, t) = \mathcal{D}(p, t) - \mathcal{S}(p, t).$$

However, we decide to approach with a simpler way for our model by analyzing only the net demand curve due to the following reasons:

1. In practice, limit orders rarely "cross" the market. Roşu (2009) states that such limit orders are irrational in equilibrium models.

2. There is a numerical instability when we model both $\mathcal{D}(p,t)$ and $\mathcal{S}(p,t)$.

Assumption 3.1.1. *There is a continuum of atomistic buyers and sellers who trade on the market. The resulting net demand curve Q is twice differentiable in price p and continuous in t. We assume that*

$$\left.\frac{\partial Q}{\partial p}\right|_{p=0} = \left.\frac{\partial Q}{\partial p}\right|_{p=S} = 0,$$
$$\left.\frac{\partial Q}{\partial p}\right|_{p} < 0 \text{ uniformly for } 0 < p < S, \text{ and } \left.\frac{\partial Q}{\partial p}\right|_{p} \leq \epsilon < 0 \;\; \mathbb{P} - a.s.$$

Remark 3.1.1. The net demand curve Q is decreasing in p.

Assuming the net demand curve $Q(p,t)$ satisfies Assumption 3.1.1, we state the stochastic differential equation (SDE) of $Q(p,t)$, for $0 \leq p \leq S$

$$\begin{aligned} dQ(p,t) &= \mu_Q(p,t)dt + \sigma_Q(p,t)\int_{s=0}^{S} b_Q(p,s,t)W(ds,dt) \\ Q(p,0) &= Q_0(p) \\ \text{where} &: \mu_Q, \sigma_Q, \text{ and } b_Q \text{ are } \mathcal{F}_t\text{-adapted}. \end{aligned} \tag{3.2}$$

We assume that the solution is unique, and uniformly bounded in p and t almost surely. Furthermore, for $\forall p$, the process $Q(p,.)$ is a semi-martingale. We also enforce

$$\int_{s=0}^{S} b_Q^2(p,s,t)ds = 1, \quad \forall p,t.$$

Definition 3.1.4. *A (trading) strategy* $\theta = (\theta(t))_{t\in[0,T]}$ *is a semi-martingale that represents a number of shares held by the investor at each point in time. If the strategy is self-financing (see Bank and Baum, 2004), the process* β^θ *representing the value of the cash account is uniquely defined.*

Definition 3.1.5. *For every real-valued* x *the inverse process* $P(x,t)$ *satisfies*

$$Q(P(x,t),t) = x. \tag{3.3}$$

The process $P(x,t)$ *is undefined when (3.3) does not admit a solution.*

Definition 3.1.6 (see (3.1) in Bank and Baum, 2004)**.** *The asymptotic liquidation proceeds* $L(\vartheta,t)$ *are defined as:*

$$L(\vartheta,t) = \int_0^\vartheta P(x,t)dx.$$

Definition 3.1.7. *The real wealth process achieved by a self-financing trading strategy* θ *is given by*

$$V^\theta(t) = \beta^\theta(t) + L(\theta(t),t).$$

Lemma 3.1.1 (see (3.2) in Bank and Baum, 2004)**.** *For any self-financing semi-martingale strategy* θ*, the dynamics of the real wealth process* V^θ *are given by*

$$\begin{aligned} V^\theta(t) - V^\theta(0_-) = &\int_0^t L(\theta(u_-),du) - \frac{1}{2}\int_0^t P'(\theta(u_-),u)d[\theta,\theta]^c_u \\ &- \sum_{0\le u\le t}\int_{\theta(u_-)}^{\theta(u)} \{P(\theta(u),u) - P(x,u)\}dx. \end{aligned} \tag{3.4}$$

Definition 3.1.8. *An arbitrage (strategy) is a self-financing trading strategy* θ *such that* $V^\theta(0_-) = 0$ *and*

$$\begin{aligned} \mathbb{P}(V^\theta(t) > 0) &> 0, \\ \mathbb{P}(V^\theta(t) > 0) &\geq 0. \end{aligned}$$

Theorem 3.1.1 (see (2.1) in German and Schellhorn, 2012)**.** *Suppose in addition to our standing assumptions that*

C1) for self-financing strategies involving only immediate orders, Jarrow (1994)'s discrete-time conditions for absence of market manipulation strategy hold,

C2) no arbitrage strategy involves wait orders,

C3) the volatility $\sigma_Q(p,t)$ is bounded away from zero, uniformly in p,

C4) there is no path such that $Q(S,t) \geq 0$ or $Q(0,t) \leq 0$.

Then

F1) there exists at least one martingale measure $\mathbb{Q}$ for $\int L(\vartheta, dt)$,

F2) there is no arbitrage strategy,

F3) the clearing price $\pi(t)$ is continuous,

F4) any such measure $\mathbb{Q}$ is also a martingale measure for $\pi(t)$.

3.2 A Market with Atomistic Traders and a Large Trader

In the previous section 3.1, we assume that there are only atomistic traders in the market. However, this assumption is quite unreal in a practical scenario. In fact, there are atomistic and large traders involving in the trading process on the stock exchange. Therefore, we modify our model to be more general with atomistic traders and a large trader.[1]

Definition 3.2.1. *The net demand curves of a large (atomistic) trader Q_L (Q_A) is a function $[0,S] \times \mathbb{R}^+ \times \Omega \rightarrow \mathbb{R}$ whose value $Q_L(p,t,\omega)$ ($Q_A(p,t,\omega)$) is equal to the difference between the quantity of shares* ***submitted*** *for purchase and the quantity of shares* ***submitted*** *for sale at price p at time t. For each p the stochastic processes $Q_L(.,t,,)$ and $Q_A(.,t,.)$ are $\mathcal{F}_t$-adapted semi-martingales. As before the net demand of the atomistic traders satisfies*

[1] For simplicity, we model a market with atomistic traders and only one large trader. If there are many large traders, our results are unchanged as long as the large traders conspire in a collusion agreement. If they do not conspire, the analysis will be much more complex since the behavior of each large trade depends on his/her expectations of the behavior of other large traders.

$$
\begin{aligned}
dQ_A(p,t) &= \mu_{Q_A}(p,t)dt + \sigma_{Q_A}(p,t)\int_{s=0}^{S} b_{Q_A}(p,s,t)W(ds,dt) \text{ for } 0 \le p \le S, \qquad (3.5)\\
Q_A(p,0) &= Q_{A,0}(p) \text{ for } 0 \le p \le S.
\end{aligned}
$$

Definition 3.2.2. *For every real-valued x the process $P_A(x,t)$ satisfies*

$$Q_A(P_A(x,t),t) = x. \qquad (3.6)$$

Definition 3.2.3. *The asymptotic liquidation proceeds of the large trader $L_L(\vartheta,t)$ are defined by*

$$L_L(\vartheta,t) = \int_0^{\vartheta} P_A(x,t)dx.$$

Assumption 3.2.1. *Both Q_L and Q_A are twice differentiable in p. Only Q_A is assumed to be continuous in t.*

Remark 3.2.1. The (total) net demand curve satisfies

$$Q = Q_L + Q_A.$$

Assumption 3.2.2. *For simplicity, we assume*

$$Q(0,t) > 0.$$

Assumption 3.2.3. *For each $p \ge \pi(t)$, the function $Q_L(p,t)$ is continuous in time.*

Lemma 3.2.1. *Let*

$$
\begin{aligned}
\mu_P(x,t) &= -\frac{\mu_{Q_A}(P_A(x,t),t) + \frac{1}{2}\frac{\partial^2 Q_A}{\partial p^2}(P_A(x,t),t)\sigma_P^2(x,t) + \frac{\partial \sigma_{Q_A}}{\partial p}(P_A(x,t),t)\sigma_P(x,t)}{\frac{\partial Q_A}{\partial p}(P_A(x,t),t)},\\
\sigma_P(x,t) &= -\frac{\sigma_{Q_A}(P_A(x,t),t)}{\frac{\partial Q_A}{\partial p}(P_A(x,t),t)}, \qquad b_P(x,s,t) = b_{Q_A}(P_A(x,t)),s,t),
\end{aligned}
$$

Then P_A is a semi-martingale and satisfies

$$dP_A(x,t) = \mu_P(x,t)dt + \sigma_P(x,t)\int_0^S b_P(x,s,t)W(ds,dt). \qquad (3.7)$$

Proof. By Definition 3.2.2, we can write

$$Q_A(P_A(x,t),t) = x. \tag{3.8}$$

We suppose (3.7) holds and apply the Ito-Wentzell formula (see Krylov, 2009) to (3.8) with given (3.5) to get

$$\begin{aligned}
dQ_A(P_A(x,t),t) &= \mu_{Q_A}(P_A(x,t),t)dt + \sigma_{Q_A}(P_A(x,t),t)\int_{s=0}^{S} b_{Q_A}(P_A(x,t),s,t)W(ds,dt) \\
&\quad + \frac{\partial Q_A}{\partial p}(P_A(x,t),t)\mu_P(x,t)dt + \frac{1}{2}\frac{\partial^2 Q_A}{\partial p^2}(P_A(x,t),t)\sigma_P^2(x,t)dt \\
&\quad + \frac{\partial Q_A}{\partial p}(P_A(x,t),t)\sigma_P(x,t)\int_{s=0}^{S} b_P(x,s,t)W(ds,dt) \\
&\quad + \frac{\partial \sigma_{Q_A}}{\partial p}(P_A(x,t),t)\sigma_P(x,t)dt. \\
&= (\mu_{Q_A}(P_A(x,t),t) + \frac{\partial Q_A}{\partial p}(P_A(x,t),t)\mu_P(x,t) \\
&\qquad + \frac{1}{2}\frac{\partial^2 Q_A}{\partial p^2}(P_A(x,t),t)\sigma_P^2(x,t) + \frac{\partial \sigma_{Q_A}}{\partial p}(P_A(x,t),t)\sigma_P(x,t))dt \\
&\quad + \int_{s=0}^{S}(\sigma_{Q_A}(P_A(x,t),t)b_{Q_A}(P_A(x,t),s,t) \\
&\qquad\qquad + \frac{\partial Q_A}{\partial p}(P_A(x,t),t)\sigma_P(x,t)b_P(x,s,t))W(ds,dt).
\end{aligned}$$

We then equate both the drift and volatility terms to zero.

$$\begin{aligned}
&\mu_{Q_A}(P_A(x,t),t) + \frac{\partial Q_A}{\partial p}(P_A(x,t),t)\mu_P(x,t) \\
&\qquad + \frac{1}{2}\frac{\partial^2 Q_A}{\partial p^2}(P_A(x,t),t)\sigma_P^2(x,t) + \frac{\partial \sigma_{Q_A}}{\partial p}(P_A(x,t),t)\sigma_P(x,t) = 0, \\
&\sigma_{Q_A}(P_A(x,t),t)b_{Q_A}(P_A(x,t),s,t) + \frac{\partial Q_A}{\partial p}(P_A(x,t),t)\sigma_P(x,t)b_P(x,s,t) = 0.
\end{aligned}$$

From the first equation, we get

$$\mu_P(x,t) = -\frac{\mu_{Q_A}(P_A(x,t),t) + \frac{1}{2}\frac{\partial^2 Q_A}{\partial p^2}(P_A(x,t),t)\sigma_P^2(x,t) + \frac{\partial \sigma_{Q_A}}{\partial p}(P_A(x,t),t)\sigma_P(x,t)}{\frac{\partial Q_A}{\partial p}(P_A(x,t),t)}.$$

From the second equation, we get

$$\sigma_P(x,t) = -\left(\frac{\sigma_{Q_A}(P_A(x,t),t)}{\frac{\partial Q_A}{\partial p}(P_A(x,t),t)}\right)\left(\frac{b_{Q_A}(P_A(x,t)),s,t)}{b_P(x,s,t)}\right).$$

Hence, the second equation holds if

$$\sigma_P(x,t) = -\frac{\sigma_{Q_A}(P_A(x,t),t)}{\frac{\partial Q_A}{\partial p}(P_A(x,t),t)}, \quad \text{and} \quad b_P(x,s,t) = b_{Q_A}(P_A(x,t)),s,t).$$

$\square$

Theorem 3.2.1 (see (2.2) in German and Schellhorn, 2012)**.** *Suppose in addition to the standing assumptions that*

C1) for self-financing strategies involving only immediate orders, Jarrow (1994)'s discrete time conditions for absence of market manipulation strategy hold,

C2) no arbitrage strategy involves wait orders,

C3) the volatility $\sigma_{Q_A}(p,t)$ is bounded away from zero, uniformly in p,

C4) there is no path such that $Q(S,t) \geq 0$ or $Q(0,t) \leq 0$.

Then

F1) there exists at least one martingale measure $\mathbb{Q}$ for $\int L_L(\vartheta, dt)$,

F2) there is no arbitrage strategy,

F3) the net demand curve Q is continuous in t,

F4) the clearing price $\pi(t)$ is continuous,

F5) any such measure $\mathbb{Q}$ is also a martingale measure for $\pi(t)$.

3.3 The Clearing Price Process under the Risk-Neutral Measure

Assumption 3.3.1. *To avoid repetition, we assume that there is no path such that $Q(S,t) \geq 0$ or $Q(0,t) \leq 0$.*

Definition 3.3.1. *The market price of risk* λ *is a function* $[0, S] \times \mathbb{R}^+ \times \Omega \to \mathbb{R}$*. The market price of risk process* $\lambda(s, ., .)$ *is an* $\mathcal{F}_t$*-adapted semi-martingale for every* $s \in [0, S]$*. We define the* $\mathbb{Q}$*-measure as a measure such that the process* $W^{\mathbb{Q}}$ *is a Brownian sheet, where:*

$$W^{\mathbb{Q}}(ds, dt) = W(ds, dt) + \lambda(s, t)dt. \tag{3.9}$$

We can then define the clearing price process as

$$\begin{aligned} d\pi(t) &= \sigma_\pi(t) \int_s b_\pi(s,t) W^{\mathbb{Q}}(ds, dt), \\ &\textit{with} \quad \int_s b_\pi(s,t)^2 ds = 1. \end{aligned} \tag{3.10}$$

From Assumption 3.1.1, $Q(p, t)$ must be strictly decreasing in p. Thus, we slightly modify the definition of $Q(p, t)$ into:

$$Q(p, t) = Q(p, 0) - \int_0^p q(y, t) dy,$$

where $Q(0, t)$ and $q(p, t)$ (for $0 < p \leq S$) are strictly positive processes uniformly bounded from zero $\mathbb{P}$-a.s. with

$$dQ(0,t) = \mu_Q(0,t)dt - \sigma_q(0,t) \int_s b_q(0,s,t) W(ds, dt), \quad \textit{with} \;\; Q(0,0) = Q_0(0) \tag{3.11}$$

$$dq(p,t) = \mu_q(p,t)dt + \sigma_q(p,t) \int_s b_q(p,s,t) W(ds, dt), \quad \textit{with} \;\; q(p,0) = Q_0(p) \tag{3.12}$$

$$\textit{for } 0 < p \leq S, \quad q(0,t) = 0,$$

$$\text{where} : \mu_Q, \mu_q, \sigma_q, b_q \text{ are } \mathcal{F}_t\text{-adapted}.$$

We assume that the solution to (3.11) and (3.12) exists, unique, and uniformly bounded in p and t almost surely. We also require

$$\int_{s=0}^S b_q^2(p, s, t) ds = 1, \quad \forall p, t.$$

Remark 3.3.1. The process q is a density of orders. By definition

$$\begin{aligned} q(p)dp \; &= \text{quantity of shares available for purchase with limit price in } [p, p+dp] \\ &\quad + \text{quantity of shares available for sale with limit price in } [p, p+dp]. \end{aligned}$$

Remark 3.3.2. Since we assume $Q(0,t)$ is twice-differentiable in p, we must make sure that $q(p,t)$ is differentiable in p, for the process $Q(p,t)$ to be twice-differentiable in p. This occurs if, for instance,

$$dq(p,t) = \int_{s=0}^{p} (p-s) W(ds, dt).$$

We now define the following processes:

$$\begin{aligned} C(\pi, t) &= -\sigma_\pi(t) \left(\frac{\partial}{\partial p} \left(\sigma_q(0,t) \int_s b_q(0,s,t) b_\pi(s,t) ds \right) + \sigma_q(\pi(t), t) \int_s b_q(\pi, s, t) b_\pi(s,t) ds \right), \\ b(\pi, t) &= -\mu_Q(0,t) + \int_0^\pi \mu_q(p,t) dp dt + \frac{1}{2} \frac{\partial q}{\partial p}(\pi, t)(\sigma_\pi(t))^2 - C(\pi, t), \\ \Sigma(\pi, s, t) &= \int_0^\pi \sigma_q(p,t) b_q(p,s,t) ds. \end{aligned}$$

Definition 3.3.2. *The market price of risk equations are:*

$$\int_{s=0}^{S} \Sigma(\pi, s, t) \lambda(s,t) ds = b(\pi, t), \;\; \textit{for } 0 \le \pi \le S.$$

Theorem 3.3.1 (see (3.1) in German and Schellhorn, 2012)**.** *Suppose that the previous assumptions hold true. In addition, suppose that the market price of risk equations have a unique solution. Then there is no arbitrage.*

Proof. Applying Ito-Wentzell formula, we can write

$$\begin{aligned} dQ(p,t) = \mu_Q(0,t)dt - \int_{0+}^{\pi} \mu_q(p,t) dp dt - \int_0^\pi \sigma_q(p,t) \int_s b_q(p,s,t) W(ds,dt) dp \\ - q(\pi,t)\sigma_\pi(t) \int_s b_\pi(s,t) W^{\mathbb{Q}}(ds,dt) - \frac{1}{2}\frac{\partial q}{\partial p}(\pi,t)(\sigma_\pi(t))^2 dt + C(\pi,t)dt \qquad (3.13) \end{aligned}$$

$$= \mu_Q(0,t)dt - \int_{0+}^{\pi} \mu_q(p,t)dpdt - \int_0^{\pi} \sigma_q(p,t) \int_s b_q(p,s,t)(W^{\mathbb{Q}}(ds,dt) - \lambda(s,t))dp$$
$$- q(\pi,t)\sigma_\pi(t) \int_s b_\pi(s,t)W^{\mathbb{Q}}(ds,dt) - \frac{1}{2}\frac{\partial q}{\partial p}(\pi,t)(\sigma_\pi(t))^2 dt + C(\pi,t)dt.$$

A market clears if $Q(p,t) = 0$, or equivalently, $dQ(p,t) = 0$; thus we set (3.13) to be zero. We equate the volatility terms to zero and find that

$$\sigma_\pi(t)b_\pi(s,t) = -\frac{\int_0^{\pi} \sigma_q(p,t)b_q(p,s,t)dp}{q(\pi,t)}. \tag{3.14}$$

Likewise, we equate the drift terms to zero.

$$\int_s \int_0^{\pi} \sigma_q(p,t)b_q(p,s,t)dp\lambda(s,t)ds =$$
$$- \mu_Q(0,t) + \int_{0+}^{\pi} \mu_q(p,t)dpdt + \frac{1}{2}\frac{\partial q}{\partial p}(\pi,t)(\sigma_\pi(t))^2 - C(\pi,t).$$

Since the above equation must hold for any value of $\pi(t)$, then the market price equations must be satisfied. Thus, there exists a unique measure $\mathbb{Q}$ such that $\pi(t)$ is a martingale. By Theorems 3.1.1 and 3.2.1, there exists a non-empty set $\mathcal{Q}$ of martingale measures for $\int L(\vartheta, dt)$. Besides, any measure $\tilde{\mathbb{Q}} \in \mathcal{Q}$ must be a martingale measure for $\pi(t)$. Uniqueness of $\mathbb{Q}$ ensures that $\mathcal{Q} = \{\mathbb{Q}\}$, thus $\mathbb{Q}$ is a martingale measure for $\int L(\vartheta, dt)$. Therefore Theorems 3.1.1 and 3.2.1 imply no arbitrage.

□

Remark 3.3.3. Integrating (3.14) and using (3.10), we see that

$$\sigma_\pi(t) = \frac{\left(\int_0^S (\int_0^{\pi} \sigma_q(p,t)b_q(p,s,t))^2 dpds\right)^{1/2}}{q(\pi,t)}. \tag{3.15}$$

The denominator of (3.15) shows that the more orders receive at the clearing price $\pi(t)$, the price volatility will likely decrease. In other word, as expected the price volatility is inversely proportional to the trading volume. The effect of the numerator is complicated to analyze, and shows that the volatility of the whole net demand curve affects the volatility of the clearing price.

In the next chapter, we will present our semi-relative no-arbitrage model of liquidity and the implementation of our model for the empirical analysis. We also show how we handle a massive NYSE Arcabook LOB data and examine them in our simulation to approximate the price and calculate the implied volatility for each European option under certain conditions.

Chapter 4

Empirical Analysis

In this chapter, we introduce a semi-relative no-arbitrage model of liquidity, developed from our theory in Chapter 3. We model the absolute quantities $Q(p, \pi(t), t)$ and $q(p, \pi(t), t)$ at the limit price p when the clearing price is $\pi(t)$, instead of the relative quantities $\widetilde{Q}(p - \pi(t), t)$ and $\widetilde{q}(p - \pi(t), t)$.[1] We decide to model this way since the absolute quantities at the limit price p when the clearing price is $\pi(t)$ are better than the relative quantities in term of keeping the clearing prices bounded when we simulate the prices.

4.1 A Semi-Relative No-Arbitrage Model of Liquidity

Let us define p_{min} and p_{max} as the minimum and maximum limit order prices respectively during the estimation period.

4.1.1 The Net Demand Curve

We define $Q(p, \pi(t), t)$ as the net demand curve at price p when the clearing price is $\pi(t)$. Then, we can model the net demand curve as

$$Q(p, \pi(t), t) = Q(p_{\text{min}}, \pi(t), t) - \int_{s=p_{\text{min}}}^{p} q(s, \pi(t), t) ds, \qquad \forall p \in [p_{\text{min}}, p_{\text{max}}] \tag{4.1}$$

[1] These relative quantities were presented in the no-arbitrage model of liquidity with relative prices in German and Schellhorn (2012).

In our model, the logarithm of the order flow quantities follow Ornstein-Uhlenbeck processes, which ensures the stationarity as well as the positivity of the order flow quantities. We can define

$$
\begin{aligned}
d\log Q(p_{\min},\pi(t),t) &= \mu_Q(p_{\min},\pi(t),t)dt + \sigma_Q(p_{\min},\pi(t),t)\int_s b_Q(p,\pi(t),s,t)W(ds,dt),\\
\text{where}: \mu_Q(p_{\min},\pi(t),t) &= -a^{rel}_{\log Q}\, Q(p_{\min},\pi(t),t)\left(\log Q(p_{\min},\pi(t),t) - \log \hat{Q}(p_{\min},\pi(t),t)\right),\\
\sigma_Q(p_{\min},\pi(t),t) &= \sigma^{rel}_{\log Q}\, Q(p_{\min},\pi(t),t). \qquad (4.2)
\end{aligned}
$$

From Definition 3.3.1, we get

$$W(ds,dt) = W^{\mathbb{Q}}(ds,dt) - \lambda(s,t)dt.$$

Therefore, we can write (4.2) under the risk-neutral $\mathbb{Q}$-measure as

$$
\begin{aligned}
d\log Q(p_{\min},\pi(t),t) &= \mu^{\mathbb{Q}}_Q(p_{\min},\pi(t),t)dt + \sigma_Q(p_{\min},\pi(t),t)\int_s b_Q(p,\pi(t),s,t)W^{\mathbb{Q}}(ds,dt),\\
\text{where}: \mu^{\mathbb{Q}}_Q(p_{\min},\pi(t),t) &= \mu_Q(p_{\min},\pi(t),t) - \sigma_Q(p_{\min},\pi(t),t)\int_s b_Q(p,\pi(t),s,t)\lambda_Q ds. \qquad (4.3)
\end{aligned}
$$

We also define

$$d\log q(p,\pi(t),t) = \mu_q(p,\pi(t),t)dt + \sigma_q(p,\pi(t),t)\int_s b_q(p,\pi(t),s,t)W(ds,dt). \qquad (4.4)$$

$$
\mu_q(p,\pi(t),t) = \begin{cases}
q(p,\pi(t),t)(-a_q(K\Delta p)(\log q(p,\pi(t),t) - A_{\hat{q}_+}(p,\pi(t),t))) \\
\qquad -\frac{1}{2}\sigma_q(p,\pi(t),t)^2\, q(p,\pi(t),t)^2, \quad p\in[\pi(t)+K\Delta p, p_{\max}] \\
q(p,\pi(t),t)(-a_q(p-\pi(t))(\log q(p,\pi(t),t) - \log \hat{q}(p,\pi(t),t))) \\
\qquad -\frac{1}{2}\sigma_q(p,\pi(t),t)^2\, q(p,\pi(t),t)^2, \quad p\in[\pi(t)-K\Delta p, \pi(t)+K\Delta p) \\
q(p,\pi(t),t)(-a_q(-K\Delta p)(\log q(p,\pi(t),t) - A_{\hat{q}_-}(p,\pi(t),t))) \\
\qquad -\frac{1}{2}\sigma_q(p,\pi(t),t)^2\, q(p,\pi(t),t)^2, \quad p\in[p_{\min}, \pi(t)-K\Delta p)
\end{cases}
$$

$$\text{where:} \quad A_{\hat{q}_+}(p,\pi(t),t) = \log \hat{q}(\pi(t)+K\Delta p)\frac{p_{\max}-p}{p_{\max}-\pi(t)+K\Delta p},$$
$$A_{\hat{q}_-}(p,\pi(t),t) = \log \hat{q}(\pi(t)-K\Delta p)\frac{p-p_{\min}}{\pi(t)-K\Delta p-p_{\min}}. \tag{4.5}$$

From (4.5), we observe that $A_{\hat{q}_+}$ is equal to zero at $p = p_{\max}$ and $A_{\hat{q}_-}$ is equal to zero at $p = p_{\min}$, which support our objective to decrease the tails while keeping the differentiability. Next, we compute $\lambda(s,t)$ from the market price of risk equations in Definition 3.3.2

$$\int_s \Sigma(\pi,s,t)\lambda(s,t)ds = b(\pi,t), \quad \text{where:} \quad \Sigma(\pi,s,t) = \int_0^{\pi} \sigma_q(p,t)b_q(p,s,t)ds.$$

We also define

$$\sigma_q(p,\pi(t),t) = \begin{cases} (p_{\max}-p)\ \sigma_q^{rel}(p-(\pi(t)+K\Delta p),t)\ q(p,\pi(t),t), & p\in[\pi(t)+K\Delta p, p_{\max}] \\ \sigma_q^{rel}(p-\pi(t),t)\ q(p,\pi(t),t), & p\in[\pi(t)-K\Delta p, \pi(t)+K\Delta p) \\ (p-p_{\min})\ \sigma_q^{rel}(p-(\pi(t)-K\Delta p),t)\ q(p,\pi(t),t), & p\in[p_{\min}, \pi(t)-K\Delta p) \end{cases}$$

In order to meet the constraint of $\int_s b_q(p,\pi(t),s,t)^2 ds = 1$, we standardize it by

$$b_q(p,\pi(t),s,t) = \frac{\widetilde{b}_q(p,\pi(t),s,t)}{\sqrt{\int_s \widetilde{b}_q(p,\pi(t),s,t)^2 ds}},$$

$$\text{where:} \quad \widetilde{b}_q(p,\pi(t),s,t) = \begin{cases} (p_{\max}-p)\ b_q^{rel}(p-(\pi(t)+K\Delta p),s), & p\in[\pi(t)+K\Delta p, p_{\max}] \\ b_q^{rel}(p-\pi(t),s), & p\in[\pi(t)-K\Delta p, \pi(t)+K\Delta p) \\ (p-p_{\min})\ b_q^{rel}(p-(\pi(t)-K\Delta),s), & p\in[p_{\min}, \pi(t)-K\Delta p) \end{cases}$$

Proof.

$$\int_s b_q(p,\pi(t),s,t)^2 ds = \int_s \left(\frac{\widetilde{b}_q(p,\pi(t),s,t)}{\sqrt{\int_s \widetilde{b}_q(p,\pi(t),s,t)^2 ds}}\right)^2 ds = \frac{\int_s \widetilde{b}_q(p,\pi(t),s,t)^2 ds}{\int_s \widetilde{b}_q(p,\pi(t),s,t)^2 ds} = 1. \quad \square$$

Similarly, we can write (4.4) under the risk-neutral $\mathbb{Q}$-measure as

$$dq(p,\pi(t),t) = \mu_q^{\mathbb{Q}}(p,\pi(t),t)dt + \sigma_q(p,\pi(t),t)\int_s b_q(p,\pi(t),s,t)W^{\mathbb{Q}}(ds,dt),$$

$$\text{where}: \mu_q^{\mathbb{Q}}(p,\pi(t),t) = \mu_q(p,\pi(t),t) - \sigma_q(p,\pi(t),t)\int_s b_q(p,\pi(t),s,t)\lambda_q(s,t)ds. \tag{4.6}$$

4.1.2 The Clearing Price Process under the Risk-Neutral Measure

We define the clearing price process under the risk-neutral $\mathbb{Q}$-measure as

$$d\pi(t) = \sigma_\pi(t)\int_s b_\pi(s,t)W^{\mathbb{Q}}(ds,dt), \tag{4.7}$$

From Remark 3.3.3, we can write

$$\sigma_\pi(t) = \frac{\left(\int_0^S (\int_0^\pi \sigma_q(p,t)b_q(p,s,t))^2 dpds\right)^{1/2}}{q(\pi,t)}.$$

After obtaining $\sigma_\pi(t)$, we can calculate $b_\pi(s,t)$ from (3.14) as follow

$$b_\pi(s,t) = -\frac{\int_0^\pi \sigma_q(p,t)b_q(p,s,t)dp}{\sigma_\pi(t)q(\pi,t)}, \quad \text{where}: \int_s b_\pi(s,t)^2 ds = 1.$$

4.2 Data

We collect our data from the NYSE Arcabook limit orders for one month, particularly in April 2011. Historical NYSE Arcabook data are recorded in the flat file ASCII format, and compressed in the GZip format, including specific information of the complete LOB from NYSE, NYSE Arca, NYSE MKT, NASDAQ and the ArcaEdge platforms from 3:30 a.m. to 8:00 p.m. ET under the high speed of latencies[2] (less than 5 milliseconds). For example, each limit order contains the unique reference number, the time stamp in seconds and milliseconds, the limit price in U.S. dollars, the quantity in number of shares, and the trading type ("B": buy or "S":

[2] Latency is defined as any delay that increases between the real and desired response times in the input/output process of the network system. *Source*: http://wordnetweb.princeton.edu.

sell). Each data point also includes the details of new adding limit orders, the modification of these orders, the deleting orders (after matching, canceling or expiring), and messages from NYSE Arca real-time data feed and information from OTC, exchange-listed, ETF and OTCBB securities. The size of one-day data is approximately 14-15GB, equivalent to 250 million rows.[3]

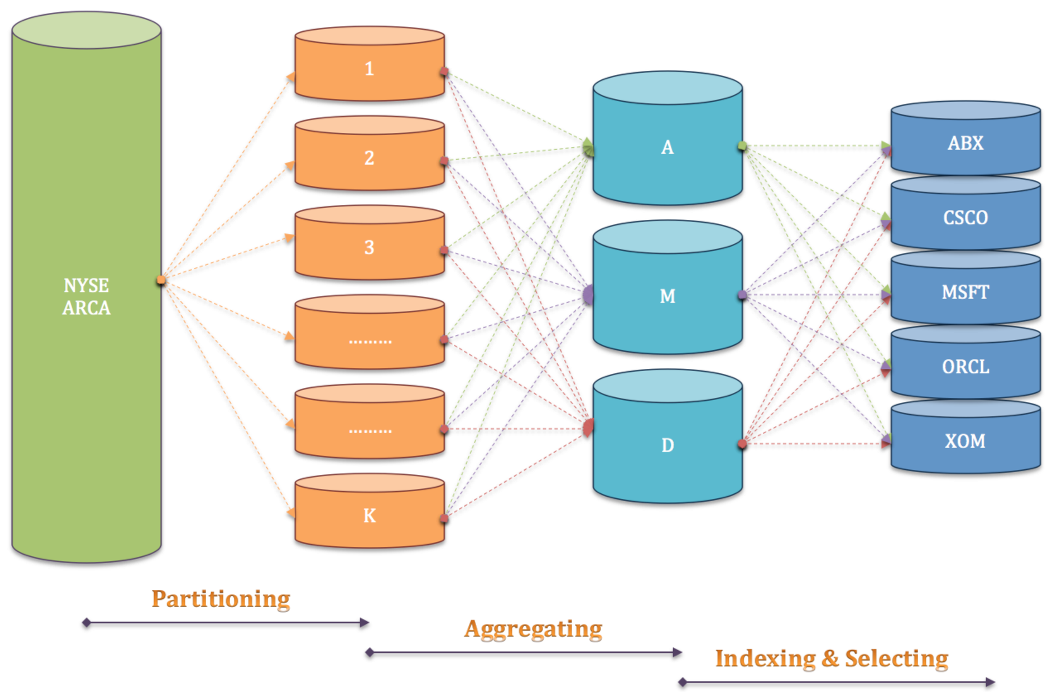


Figure 4.1: NYSE Arcabook data handling process.

In order to handle this massive data set, we develop and employ the parallelization algorithm in which, we partition the large data file into smaller files, and distribute them to multiple hosts (i.e. computers) for indexing and selecting purposes.[4] For example, we partition a 250-million-row data set for one day into 250 one-million-row data sets, then assign them into different computers for indexing and selecting types "A", "M" or "D". Next, we aggregate these files into three large subsets "A", "M", "D" showed in Figure 4.1.[5] Our final step is to

[3] Total of one-month data are approximately 300GB, equivalent to 5.25 billion rows.
[4] We automated the entire process of the parallelization algorithm in R.
[5] NYSE Arcabook FTP Client Specification, version 1.5a, June 21, 2011 specifies the message type "A" which

select stocks based on their trading tickers from these three large subsets.

To support our empirical analysis, we focus our research on the HF data during the core trading section[6] from 9:30 a.m. to 4:00 p.m. ET and select at least two different stocks from each industry as follow

Industry	**Exchange**	**Ticker**	**Firm**
Energy	NYSE	CVX	Chevron Corporation
	NYSE	XOM	Exxon Mobil Corporation
Financial Banks	NYSE	JPM	JPMorgan Chase & Co.
	NYSE	WFC	Wells Fargo & Company
Materials and Mining	NYSE	ABX	Barrick Gold Corporation
	NYSE	FCX	Freeport-McMoRan Copper & Gold Inc.
Technology	NASDAQ	CSCO	Cisco Systems, Inc.
	NASDAQ	MSFT	Microsoft Corporation
	NASDAQ	ORCL	Oracle Corporation

Table 4.1: NYSE Arcabook data selection.

4.2.1 Cancelation Methodology for Limit Orders

There are no identifier among canceled, expired or filled orders in the NYSE Arcabook data, since all of them are indicated as type "D". Therefore; to support our research[7], we propose a cancelation methodology:

1. The limit orders which do not match during the trading day will be canceled at the end of the trading day, consistent with the rule of NYSE Arca.

2. Since we can obtain the highest and lowest historical prices of each trading day from

indicates a new order added to NYSE Arca's open order with a unique identifier; the message type "M" which indicates: (1) the price of an order changes, (2) the size of an order changes or (3) an order is partially filled; the message type "D" which indicates: (1) an order is cancelled, (2) an order expires, or (3) an order is filled.

[6] *Source*: http://www.nyx.com/en/holidays-and-hours/arca.

[7] In general, canceled orders not only enhance the dynamics of the limit order flows, but also make the data more realistic.

the public information[8] and all unexecuted limit orders will be canceled at the end of the trading day, we can reasonably assume that the limit orders with their limit prices which are out of range between highest and lowest prices during the trading day will be canceled when the "D" condition of these limit orders occurred.

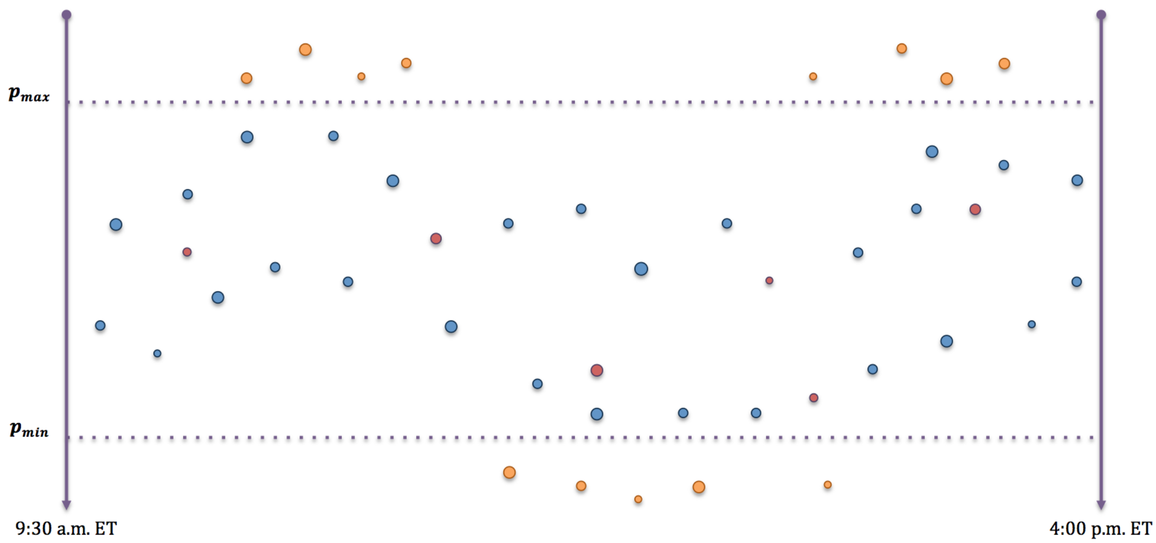


Figure 4.2: Cancelation methodology.

Furthermore, we discover that under our methodology, the cancelation rates are approximately up to 20% in Appendix A, consistent with Gould et al. (2012).

4.2.2 Selection Methodology for Clearing Prices

By dividing the core trading section from 9:30 a.m. to 4:00 p.m. ET into one-minute intervals, we obtain 390[9] time intervals.

[8] *Source*: Yahoo! Finance, Google Finance.
[9] 6.5 *active trading hours* multiplied by 60 *minutes* equals 390 *one-minute* intervals.

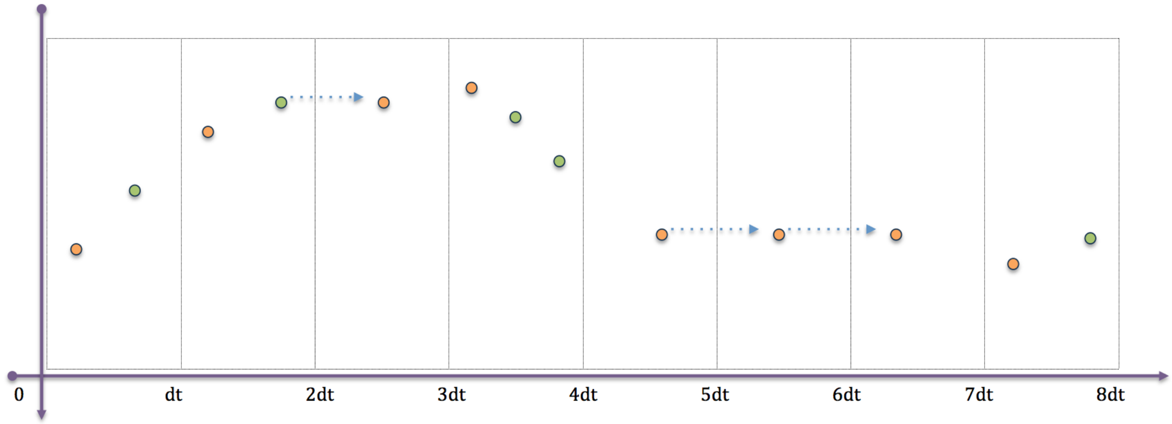


Figure 4.3: Selection methodology for the clearing prices $\pi(t)$.

To obtain the clearing price $\pi(t)$ for each time interval assuredly[10], we need to employ some certain conditions as follow:

1. Select the *first* clearing price in the current time interval as the representative $\pi(t)$ of that time interval.

2. Select the *last* clearing price in the previous time interval as the representative $\pi(t)$ of the current time interval if there is no clearing price in the current time interval.

3. Select the representative $\pi(t-1)$ as the representative $\pi(t)$ of the current time interval if there is no clearing price in both current and previous time intervals.

4. Select the *minimum* clearing price of the trading day as the representative $\pi(t)$ of the first time interval if there is no clearing price in that time interval.

[10] We prove that the selection method performs properly in Appendix B.

4.3 Simulation Methodology

In our simulation, we discretize our model in three dimensions including the limit order price p, the factor s with $s \in [p_{\text{min}} + \Delta p, p_{\text{max}}]$, and the trading time t.

From (4.3), we can write the recursion formula of log $Q(p_{\text{min}}, \pi(t), t)$ as follow[11]

$$\begin{aligned}\log Q(p_{\text{min}}, \pi(t+\Delta t), t+\Delta t) &= \log Q(p_{\text{min}}, \pi(t), t) + \mu_Q^{\mathbb{Q}}(p_{\text{min}}, \pi(t), t) + \sigma_Q(p_{\text{min}}, \pi(t), t) \\ &\quad \sum_{s=p_{\text{min}}+\Delta p}^{p_{\text{max}}} b_Q(p, \pi(t), s, t) Z(s,t)\sqrt{\Delta t}\sqrt{\Delta p}, \\ \text{where}: \mu_Q^{\mathbb{Q}}(p_{\text{min}}, \pi(t), t) &= \mu_Q(p_{\text{min}}, \pi(t), t) - \sigma_Q(p_{\text{min}}, \pi(t), t) \\ &\quad \sum_{s=p_{\text{min}}+\Delta p}^{p_{\text{max}}} b_Q(p, \pi(t), s, t)\lambda_Q \Delta p. \end{aligned} \tag{4.8}$$

From (4.6), similarly we can write the recursion formula of log $q(p, \pi(t), t)$

$$\begin{aligned}\log q(p, \pi(t+\Delta t), t+\Delta t) &= \log q(p, \pi(t), t) + \mu_q^{\mathbb{Q}}(p, \pi(t), t) + \sigma_q(p, \pi(t), t) \\ &\quad \sum_{s=p_{\text{min}}+\Delta p}^{p_{\text{max}}} b_q(p, \pi(t), s, t) Z(s,t)\sqrt{\Delta t}\sqrt{\Delta p}, \\ \text{where}: \mu_q^{\mathbb{Q}}(p, \pi(t), t) &= \mu_q(p, \pi(t), t) - \sigma_q(p, \pi(t), t) \\ &\quad \sum_{s=p_{\text{min}}+\Delta p}^{p_{\text{max}}} b_q(p, \pi(t), s, t)\lambda_q(s,t) \Delta p. \end{aligned} \tag{4.9}$$

From (4.7), we can write the recursion formula of the clearing price process $\pi(t)$

$$\pi(t+\Delta t) = \pi(t) + \sigma_\pi(t) \sum_{s=p_{\text{min}}+\Delta p}^{p_{\text{max}}} b_\pi(s,t) Z(s,t)\sqrt{\Delta t}\sqrt{\Delta p} \tag{4.10}$$

[11] *Note*: $Z(s,t)$ is a vector of random i.i.d. standard normal variables ($\mu_Z = 0$, $\sigma_Z = 1$).

Simulation steps:

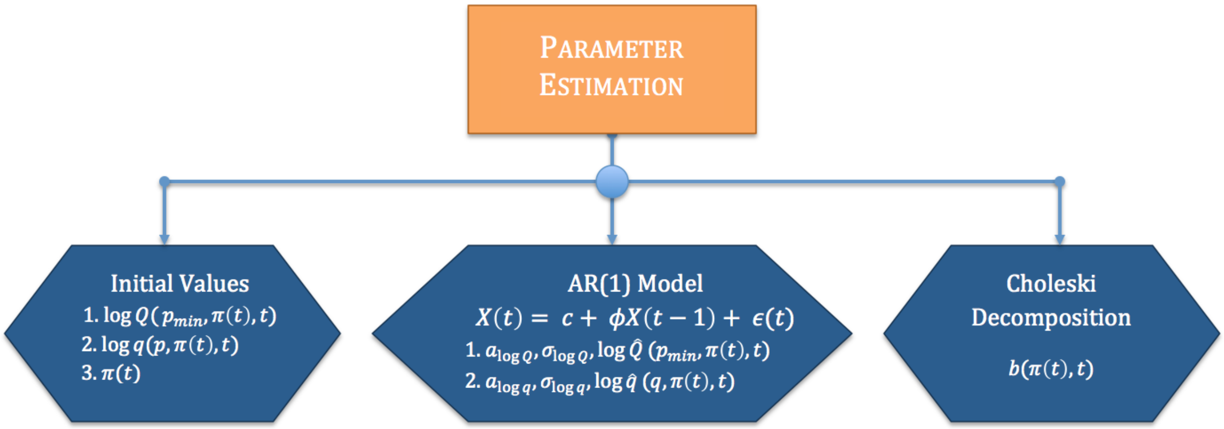


Figure 4.4: Parameter Estimation.

Step 1: We calculate $\log Q(p_{\min}, \pi(t), t)$, $\log q(p, \pi(t), t)$ and $\pi(t)$ at the end of the trading day on April 1, 2011 for the initial values of our simulation.

Step 2: We implement the AR(1) model[12] for $\log Q(p_{\min}, \pi(t), t)$ and $\log q(p, \pi(t), t)$ based on the trading data of April 4, 2011[13]

$$X(t) = c + \phi X(t-1) + \epsilon(t), \quad \text{where} : \epsilon(t) \sim N(\mu_\epsilon, \sigma_\epsilon).$$

We then estimate for[14]

- $\log \hat{Q}(p_{\min}, \pi(t), t)$: speed mean reversion $a^{rel}_{\log Q}$ as $(1 - \phi_Q)$, volatility $\sigma^{rel}_{\log Q}$ as $\sigma_{\mathrm{AR}(1)_Q}$, and long-term value $\log \hat{Q}^{rel}$ as c_Q.
- $\log q(p, \pi(t), t)$: speed mean reversion $a^{rel}_{\log q}(k\Delta p)$ as $(1 - \phi_q)$, volatility $\sigma^{rel}_{\log q}(k\Delta p)$ as $\sigma_{\mathrm{AR}(1)_q}$, and long-term value $\log \hat{q}^{rel}(k\Delta p)$ as c_q.

[12] For simplicity, we choose the simplest autoregressive model of order (1) based on the result of the autocorrelation function (ACF) in R.

[13] We have to skip two days (April 2, 2011 and April 3, 2011) for those are on weekends.

[14] Similarly in German and Schellhorn (2012), let us define the relative price k as $k \equiv p - \pi(t)$. Therefore, in term of value, we notice that (1) $a_{\log Q}(p_{\min}, \pi(t), t)$, $\log \hat{Q}(p_{\min}, \pi(t), t)$ are equivalent to $a^{rel}_{\log Q}$ and $\log \hat{Q}^{rel}$; (2) $a_{\log q}(p, \pi(t), t)$, $\log \hat{q}(p, \pi(t), t)$ are equivalent to $a^{rel}_{\log q}(k\Delta p)$ and $\log \hat{q}^{rel}(k\Delta p)$ respectively for $k \in [-K+1, K]$.

Step 3: For N number of scenarios, we simulate the lognormal models and the clearing price process $\pi(t)$ based on the formulas (4.8), (4.9) and (4.10) with

$$\sigma_\pi(t) = \frac{\left(\sum_0^S (\sum_0^\pi \sigma_q(p,t) b_q(p,s,t))^2 \Delta p \Delta s\right)^{1/2}}{q(\pi, t)},$$

$$b_\pi(s,t) = -\frac{\sum_0^\pi \sigma_q(p,t) b_q(p,s,t) \Delta p}{\sigma_\pi(t) q(\pi,t)}, \quad \text{where} : \sum_{s=p_{\min}+\Delta p}^{p_{\max}} b_\pi(s,t)^2 \Delta p = 1.$$

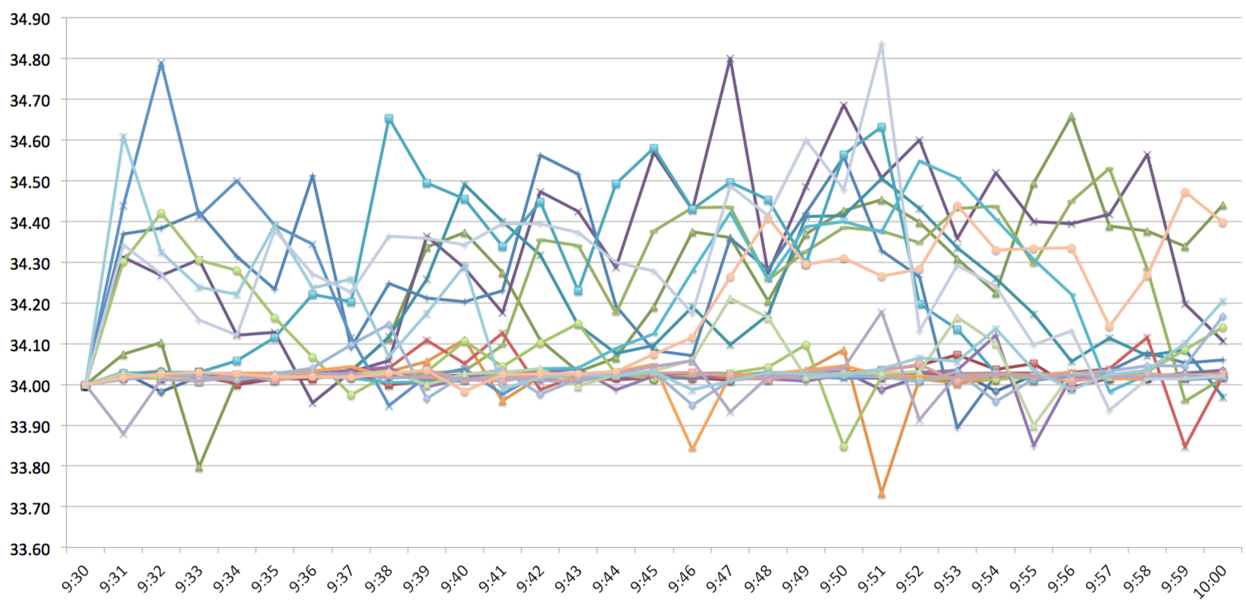


Figure 4.5: Simulated $\pi_{\text{ORCL}}(t)$ for 25 scenarios in the first 30 minutes..

We notice that during the simulation process, the simulated clearing prices $\pi(t)$ could be out of bound of $[p_{\min}, p_{\max}]$ where data are short for the next simulation step. Therefore, under the data constraint, we need to simulate the clearing prices $\pi(t)$ under certain conditions that if the simulated clearing price is lower than $p_{\min}$ or higher than $p_{\max}$, we set the clearing price for that time interval as $p_{\min}$ or $p_{\max}$ respectively for the next simulation step, but we still keep the simulated $\pi(t)$ in our output of clearing prices. For example, in the ORCL example for 25 scenarios in the first 30 minutes in Figure 4.5, we observe that our simulated clearing prices $\pi(t)$ from \$33.73 to \$34.84 are out of bound of [\$34.01, \$34.43] on April 4, 2011.

Step 4: Finally, we price a European call option approximately under the Monte Carlo simulation over N number of scenarios and compare the simulated option price with the value of the European call option from the Black-Scholes-Merton model[15] where σ_{BSM} equals to the constant standard deviation of N number of simulated clearing prices at time T.

$$C(\psi) = \frac{1}{N}\sum_{\omega=1}^{N} \text{max}(\pi(T,\omega) - \psi, 0),$$

$$\text{where :} \quad \psi = \text{a strike price of the option},$$

$$T = \text{a maturity time}.$$

For each European option, we also calculate the implied volatility and compare our simulated volatility smile with the actual one from Bloomberg.

In the next chapter, we will present our empirical results from nine stocks in four different industries, such as energy (CVX, XOM), financial banking (JPM, WFC), materials and mining (ABX, FCX) and technology (CSCO, MSFT, ORCL). In order to evaluate the plausibility of our model, we divide these results into two groups: (1) favorable results, and (2) open issues for extensive studies.

[15] In the Black-Scholes-Merton model, the value of a European call option for a non-dividend stock:

$$C(S,t) = SN(d_1) - Ke^{-r(T-t)}N(d_2)$$

$$\text{where :} \quad d_1 = \frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right)$$

$$d_2 = \frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{S}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)(T-t)\right) = d_1 - \sigma\sqrt{T-t}$$

S : *a price of the stock*, K : *a strike price of the option*,

r : *an annualized risk-free interest rate*,

σ : *a volatility of the stock's returns*,

T : *a maturity time*.

Chapter 5

Empirical Results

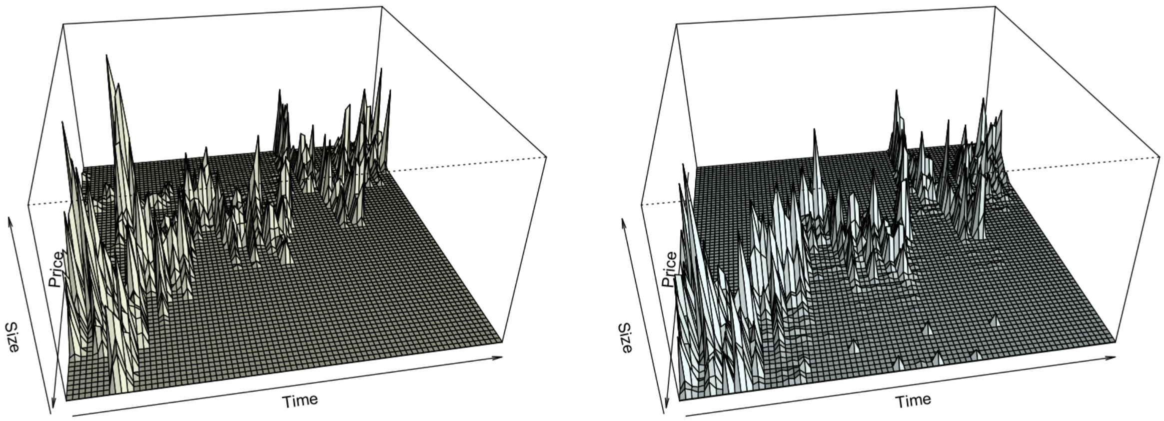
In this chapter, we specifically present how to implement our simulation methodology on each stock. We approximate the price of each European option based on each underlying stock and compare our simulated price to the value from the Black-Scholes-Merton model. We also present our simulated volatility smile and compare it to the actual one from Bloomberg, and separate these results into two groups: (1) favorable results, and (2) open issues for extensive studies in order to evaluate the plausibility of our model.

5.1 Favorable Results

5.1.1 Barrick Gold Corporation (ABX)

Let $p_{\text{ABX}_{\text{min}}}$ be the minimum price of USD 51.31 and $p_{\text{ABX}_{\text{max}}}$ be the maximum price of USD 52.05 during the core trading section from 9:30 a.m. to 4:00 p.m. ET on April 4, 2011. We partition the price range into one-cent bins and the time span into one-minute bins

$$
\begin{aligned}
p_{\text{ABX}_{\text{min}}} = \$51.31 = p_{-K} < p_{-K+1} < \cdots < p_{K-1} < p_K = p_{\text{ABX}_{\text{max}}} = \$52.05, \\
p_{K+1} - p_K = \Delta p \text{ (one cent)}, \\
t_{i+1} - t_i = \Delta t \text{ (one minute)}.
\end{aligned}
$$



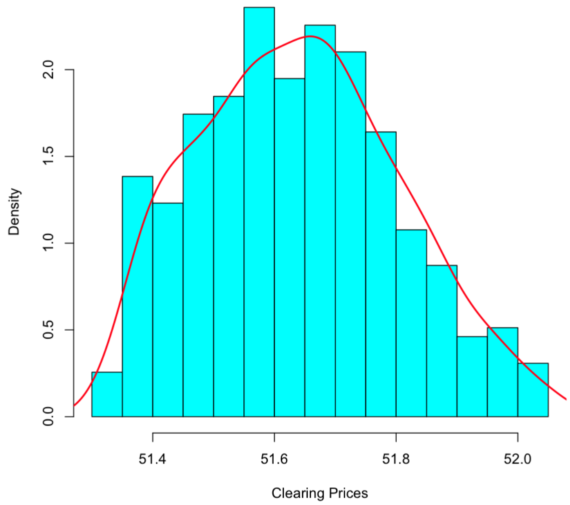
(a) Buy Limit Orders (b) Sell Limit Orders

Figure 5.1: The dynamics of Barrick Gold Corporation's LOB on April 4, 2011.

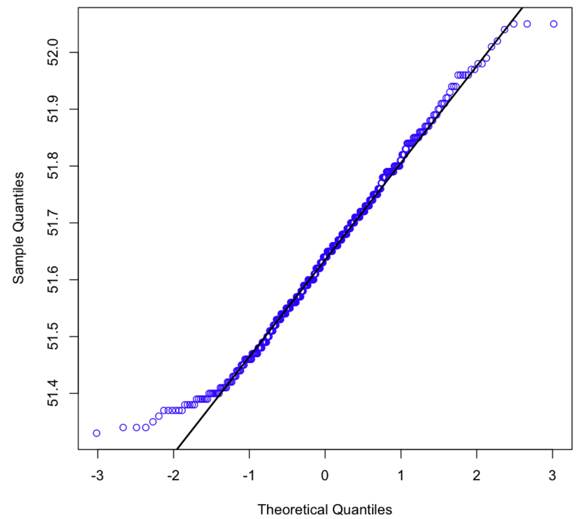
By dividing the core trading section from 9:30 a.m. to 4:00 p.m. ET into one-minute intervals, we obtain 390 time intervals and present the summary statistics of $\pi_{\text{ABX}}(t)$ on April 4, 2011 in Table 5.1 as follow

Summary Statistics			
nobs	390	Sum	20140.01
NAs	0	SE Mean	0.0083
Minimum	51.31	LCL Mean	51.6248
Maximum	52.05	UCL Mean	51.6573
1. Quartile	51.52	Variance	0.0267
3. Quartile	51.75	Stdev	0.1634
Mean	51.6411	Skewness	0.2363
Median	51.6400	Kurtosis	-0.6118

Table 5.1: Summary statistics of $\pi_{\text{ABX}}(t)$ (1 minute) on April 4, 2011.



(a) The histogram of $\pi_{\text{ABX}}(t)$



(b) Normal Q-Q plot of $\pi_{\text{ABX}}(t)$

Figure 5.2: $\pi_{\text{ABX}}(t)$ (1 minute) on April 4, 2011.

We conduct the parameter estimation of our model based on the simulation steps over the limit order data of Barrick Gold Corporation (ABX) described in Chapter 4. From the observed data, the opening clearing price of ABX on April 4, 2011 was

$$\pi_{\text{ABX}}(0) = \$51.32.$$

We then simulate the lognormal models and the clearing price process $\pi(t)$ based on the formulas (4.8), (4.9) and (4.10) respectively with

$$\sigma_\pi(t) = \frac{\left(\sum_0^S (\sum_0^\pi \sigma_q(p,t) b_q(p,s,t))^2 \Delta p \Delta s\right)^{1/2}}{q(\pi,t)},$$

$$b_\pi(s,t) = -\frac{\sum_0^\pi \sigma_q(p,t) b_q(p,s,t)\Delta p}{\sigma_\pi(t) q(\pi,t)}, \quad \text{where} : \sum_{s=p_{\min}+\Delta p}^{p_{\max}} b_\pi(s,t)^2 \Delta p = 1.$$

Table 5.2: The parameter estimation of Barrick Gold Corporation (ABX).

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log Qtilde $p_{\min}$	21.666488	0.000131	0.053365	20.291618
log qtilde(-36 , t)	14.611532	0.000674	0.095028	14.372984
log qtilde(-35 , t)	14.969530	0.005423	0.578003	10.787664
log qtilde(-34 , t)	15.365796	0.006177	0.631129	12.693542
log qtilde(-33 , t)	15.762062	0.004439	0.596245	12.332508
log qtilde(-32 , t)	15.686874	0.000023	0.578997	14.822731
log qtilde(-31 , t)	15.135407	0.001019	0.107815	13.692149
log qtilde(-30 , t)	15.644491	0.000000	0.581997	15.290412
log qtilde(-29 , t)	15.498981	0.004880	0.618303	12.262985
log qtilde(-28 , t)	15.741358	0.004362	0.644635	15.779390
log qtilde(-27 , t)	15.168592	0.000045	0.576088	15.161519
log qtilde(-26 , t)	16.510773	0.000514	0.067654	15.364727
log qtilde(-25 , t)	16.433546	0.000437	0.044014	14.861389
log qtilde(-24 , t)	16.023425	0.000860	0.062038	13.992369
log qtilde(-23 , t)	17.017893	0.000709	0.086119	14.993793
log qtilde(-22 , t)	16.768777	0.000795	0.080151	15.127205
log qtilde(-21 , t)	16.936688	0.000299	0.040225	14.475750
log qtilde(-20 , t)	16.536000	0.000706	0.100273	14.647169
log qtilde(-19 , t)	17.367656	0.000628	0.113647	15.493425
log qtilde(-18 , t)	17.400008	0.000009	0.586285	15.946711
log qtilde(-17 , t)	17.531546	0.000520	0.057701	14.909882
log qtilde(-16 , t)	17.741395	0.002875	0.572299	12.990529
log qtilde(-15 , t)	17.817423	0.000590	0.106191	14.384678
log qtilde(-14 , t)	17.857440	0.000679	0.115211	15.256224
log qtilde(-13 , t)	17.760328	0.000771	0.094671	15.339394
log qtilde(-12 , t)	18.390178	0.000664	0.105254	14.781018
log qtilde(-11 , t)	18.201616	0.000387	0.075082	15.369834
log qtilde(-10 , t)	18.277280	0.000270	0.056353	16.177634
log qtilde(-9 , t)	18.000476	0.003604	0.603660	13.317607
log qtilde(-8 , t)	18.541096	0.000280	0.068767	17.564520
log qtilde(-7 , t)	18.345194	0.000271	0.073230	17.161658
log qtilde(-6 , t)	18.344671	0.000198	0.052732	17.182301

Continued on next page

Table 5.2 -- *Continued from previous page*

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log qtilde(-5 , t)	19.055585	0.000218	0.064850	17.663597
log qtilde(-4 , t)	18.718420	0.000178	0.057220	17.664019
log qtilde(-3 , t)	18.844917	0.000194	0.068881	17.746271
log qtilde(-2 , t)	18.858403	0.000002	0.604809	17.823426
log qtilde(-1 , t)	19.521784	0.000178	0.072216	18.051623
log qtilde(0 , t)	18.940949	0.000232	0.089195	17.192191
log qtilde(1 , t)	18.857619	0.000356	0.106258	16.440069
log qtilde(2 , t)	19.479261	0.000238	0.089136	17.658444
log qtilde(3 , t)	18.765429	0.000213	0.070157	16.476920
log qtilde(4 , t)	18.767376	0.000387	0.087883	15.544154
log qtilde(5 , t)	18.838702	0.000847	0.145184	15.271946
log qtilde(6 , t)	19.101994	0.000279	0.095131	17.746553
log qtilde(7 , t)	18.477541	0.000218	0.067431	17.085007
log qtilde(8 , t)	18.548345	0.000418	0.091841	16.740316
log qtilde(9 , t)	18.805478	0.000258	0.088835	17.596702
log qtilde(10 , t)	18.192758	0.000295	0.078281	16.472143
log qtilde(11 , t)	18.001128	0.000194	0.057462	16.466912
log qtilde(12 , t)	18.310868	0.000264	0.070510	16.167729
log qtilde(13 , t)	18.475999	0.000347	0.083607	16.450186
log qtilde(14 , t)	17.843511	0.002507	0.564261	14.187474
log qtilde(15 , t)	17.916335	0.000232	0.049943	15.519892
log qtilde(16 , t)	17.678644	0.000505	0.101438	15.159032
log qtilde(17 , t)	17.776435	0.000215	0.056530	15.954785
log qtilde(18 , t)	17.275990	0.003518	0.558781	13.094909
log qtilde(19 , t)	17.391893	0.000420	0.077610	15.011230
log qtilde(20 , t)	17.769947	0.005141	0.619041	13.222007
log qtilde(21 , t)	16.864435	0.000566	0.065946	14.486075
log qtilde(22 , t)	17.028093	0.005534	0.594576	13.601811
log qtilde(23 , t)	17.101774	0.007363	0.623889	12.821840
log qtilde(24 , t)	17.246193	0.006155	0.635907	11.992500
log qtilde(25 , t)	15.936491	0.005961	0.574075	12.706216
log qtilde(26 , t)	16.654246	0.005246	0.599633	12.874782
log qtilde(27 , t)	17.068197	0.003197	0.592514	15.608349

Continued on next page

Table 5.2 -- *Continued from previous page*

Parameter	**Initial Value**	**Speed Mean Reversion**	**Volatility**	**Long-Term Value**
log qtilde(28 , t)	15.857147	0.004278	0.576015	12.738550
log qtilde(29 , t)	16.423008	0.004650	0.597749	13.096454
log qtilde(30 , t)	16.183542	0.003229	0.545302	12.999579
log qtilde(31 , t)	16.606710	0.004085	0.589484	13.084485
log qtilde(32 , t)	16.162492	0.004639	0.599101	12.949548
log qtilde(33 , t)	16.304474	0.005323	0.598722	12.390814
log qtilde(34 , t)	16.648027	0.004510	0.594052	12.820601
log qtilde(35 , t)	15.762967	0.003820	0.561960	12.793149
log qtilde(36 , t)	15.844196	0.004581	0.590985	12.834974
log qtilde(37 , t)	15.737560	0.004039	0.570404	12.792506
Clearing Price $\pi(0)$	51.32			

By using 500 scenarios, we approximate the value of a European call option

$$C(\psi) = \frac{1}{N} \sum_{\omega=1}^{N} \text{max}(\pi(T, \omega) - \psi, 0),$$

$$\begin{aligned} \text{where :} \quad & N = 500, \\ & \psi = \text{a strike price of the option}, \\ & T = \frac{1}{365}, \text{ that is equal to one day.} \end{aligned}$$

We also compare our simulated option price with the approximated value of a European call option for the same underlying stock from the Black-Scholes-Merton (BSM) model with a constant σ_{BSM} of 24.97%. As a result, we find that our simulated prices are closed to the BSM values in Table 5.3, which proves the accuracy of our model.

Strike Price	Implied Volatility	Simulated Price	BSM Price	Abs(Sim - BSM)
50.99	24.56%	0.458468	0.462326	0.003859
51.04	22.86%	0.407527	0.428082	0.020555
51.09	21.11%	0.356664	0.395265	0.038602
51.14	19.29%	0.305825	0.363915	0.058091
51.20	17.40%	0.255094	0.334064	0.078970
51.25	15.49%	0.205192	0.305734	0.100542
51.30	13.55%	0.156687	0.278941	0.122254
51.35	12.87%	0.124216	0.253692	0.129476
51.40	13.41%	0.107653	0.229982	0.122329
51.45	13.94%	0.093348	0.207801	0.114453
51.50	14.43%	0.080867	0.187128	0.106260
51.55	14.83%	0.069371	0.167935	0.098564
51.60	15.14%	0.058810	0.150186	0.091375
51.65	15.37%	0.049227	0.133837	0.084610
51.70	15.52%	0.040544	0.118839	0.078295
51.76	15.66%	0.033108	0.105137	0.072029
51.81	15.80%	0.026908	0.092671	0.065763
51.86	15.87%	0.021448	0.081376	0.059929
51.91	15.84%	0.016540	0.071188	0.054648
51.96	15.72%	0.012310	0.062035	0.049726

Table 5.3: Simulation summary of European call options on ABX ($\sigma_{\text{BSM}} = 24.97\%$).

Furthermore, we calculate the implied volatility of the European call options for each strike price. Our at-the-money (ATM) value of $51.35 is very closed to the ATM value of $51.40 from Bloomberg. Most importantly, our simulation result in Figure 5.3a shows a volatility smile[1] which is observed in real markets in Figure 5.3b, which confirms the plausibility of our model.

[1] In finance, a volatility smile is a pattern in which the implied volatility is relatively low for at-the-money options and higher for either in-the-money or out-of-the-money options. *Source*: Hull (2005).

(a) Simulated implied volatility as a function of the strike price,

(b) Actual implied volatility as a function of the strike price from Bloomberg.

Figure 5.3: Implied volatility as a function of the strike price of ABX.

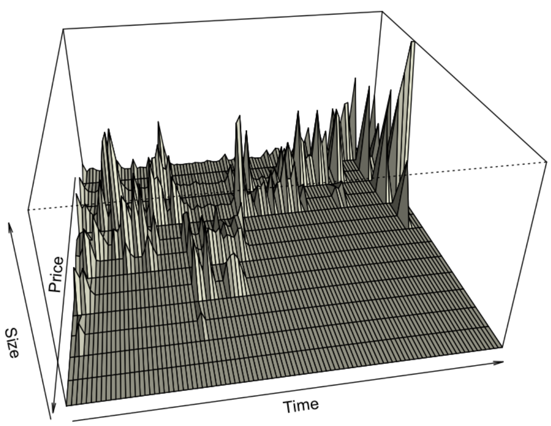
5.1.2 Cisco Systems, Inc. (CSCO)

Let $p_{\text{CSCO}_{\min}}$ be the minimum price of USD 17.00 and $p_{\text{CSCO}_{\max}}$ be the maximum price of USD 17.14 during the core trading section from 9:30 a.m. to 4:00 p.m. ET on April 4, 2011. We partition the price range into one-cent bins and the time span into one-minute bins

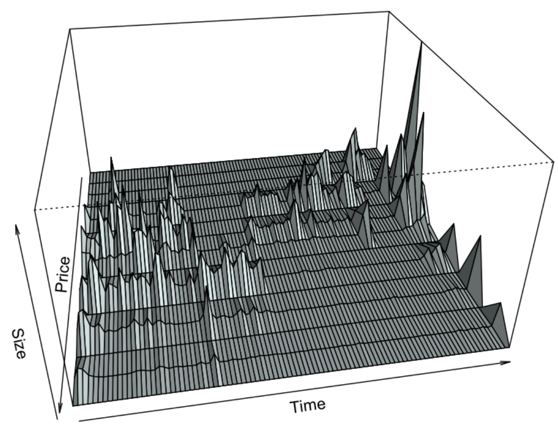
$$p_{\text{CSCO}_{\min}} = \$17.00 = p_{-K} < p_{-K+1} < \cdots < p_{K-1} < p_K = p_{\text{CSCO}_{\max}} = \$17.14,$$

$$p_{K+1} - p_K = \Delta p \text{ (one cent)},$$

$$t_{i+1} - t_i = \Delta t \text{ (one minute)}.$$



(a) Buy Limit Orders



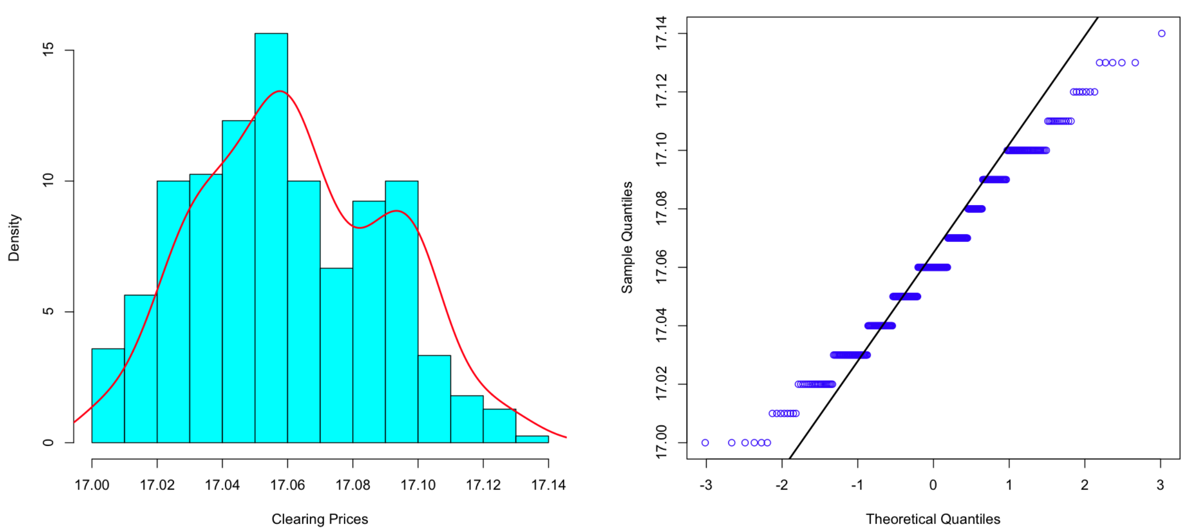
(b) Sell Limit Orders

Figure 5.4: The dynamics of Cisco Systems, Inc.'s LOB on April 4, 2011.

By dividing the core trading section from 9:30 a.m. to 4:00 p.m. ET into one-minute intervals, we obtain 390 time intervals and present the summary statistics of $\pi_{\text{CSCO}}(t)$ on April 4, 2011 in Table 5.4 as follow

Summary Statistics			
nobs	390	Sum	6654.36
NAs	0	SE Mean	0.0015
Minimum	17.00	LCL Mean	17.0596
Maximum	17.14	UCL Mean	17.0654
1. Quartile	17.04	Variance	0.0008
3. Quartile	17.09	Stdev	0.0290
Mean	17.0625	Skewness	0.1598
Median	17.0600	Kurtosis	-0.6126

Table 5.4: Summary statistics of $\pi_{\text{csco}}(t)$ (1 minute) on April 4, 2011.



(a) The histogram of $\pi_{\text{csco}}(t)$

(b) Normal Q-Q plot of $\pi_{\text{csco}}(t)$

Figure 5.5: $\pi_{\text{csco}}(t)$ (1 minute) on April 4, 2011.

Similarly, we conduct the parameter estimation of our model based on the simulation steps over the limit order data of Cisco Systems, Inc. (CSCO) described in Chapter 4. From the observed data, the opening clearing price of CSCO on April 4, 2011 was

$$\pi_{\text{csco}}(0) = \$17.03.$$

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log Qtilde $p_{\min}$	21.915265	0.000108	0.029489	20.178658
log qtilde(-6 , t)	18.934207	0.000270	0.041408	17.318796
log qtilde(-5 , t)	18.654855	0.000550	0.093933	17.159957
log qtilde(-4 , t)	19.562341	0.000547	0.106918	17.925457
log qtilde(-3 , t)	19.355660	0.000329	0.027911	17.719784
log qtilde(-2 , t)	20.030462	0.000122	0.029936	18.613246
log qtilde(-1 , t)	19.486814	0.000202	0.035705	17.874711
log qtilde(0 , t)	21.077751	0.000166	0.047992	19.120142
log qtilde(1 , t)	20.802213	0.000206	0.047965	19.214699
log qtilde(2 , t)	19.205480	0.000229	0.058813	17.475067
log qtilde(3 , t)	20.129510	0.000253	0.030598	18.759752
log qtilde(4 , t)	18.786959	0.001821	0.206094	17.472670
log qtilde(5 , t)	19.537187	0.000399	0.048883	18.428087
log qtilde(6 , t)	18.368037	0.000000	0.246592	17.549220
log qtilde(7 , t)	19.194682	0.000406	0.039625	17.824362
Clearing Price $\pi(0)$	17.03			

Table 5.5: The parameter estimation of Cisco Systems, Inc. (CSCO).

We simulate the lognormal models and the clearing price process $\pi(t)$ based on the formulas (4.8), (4.9) and (4.10) respectively, and then we approximate the value of a European call option with 500 scenarios

$$C(\psi) = \frac{1}{N}\sum_{\omega=1}^{N} \max(\pi(T,\omega) - \psi, 0),$$

$$\text{where :} \quad N = 500,$$

$$\psi = \text{a strike price of the option},$$

$$T = \frac{1}{365}, \text{ that is equal to one day.}$$

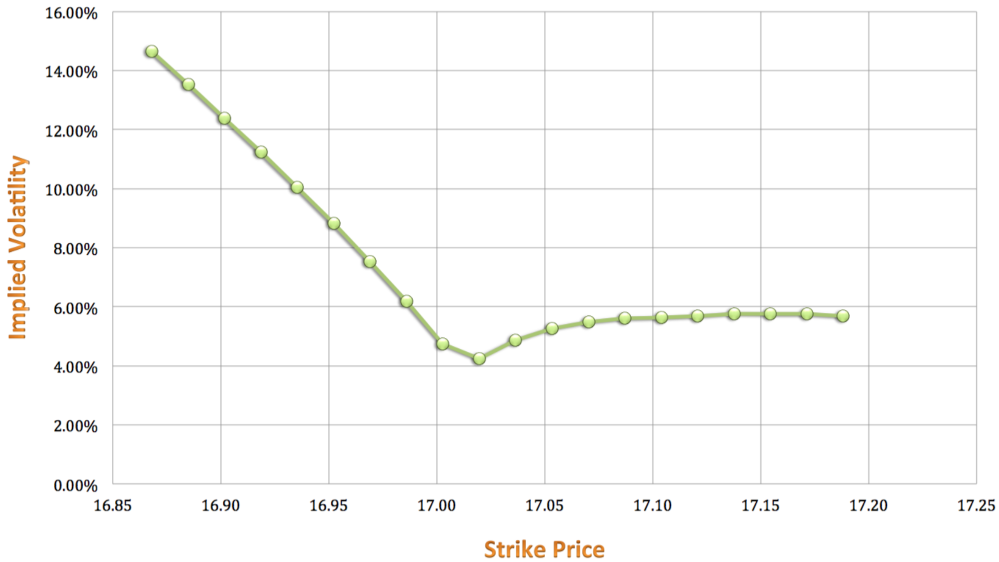
Likewise, we compare our simulated option price with the approximated value of a European call option for the same underlying stock from the Black-Scholes-Merton model with a constant σ_{BSM} of 3.56%. As a result, we find that our simulated prices are very closed to the

BSM values in Table 5.6, which once again proves the accuracy of our model.

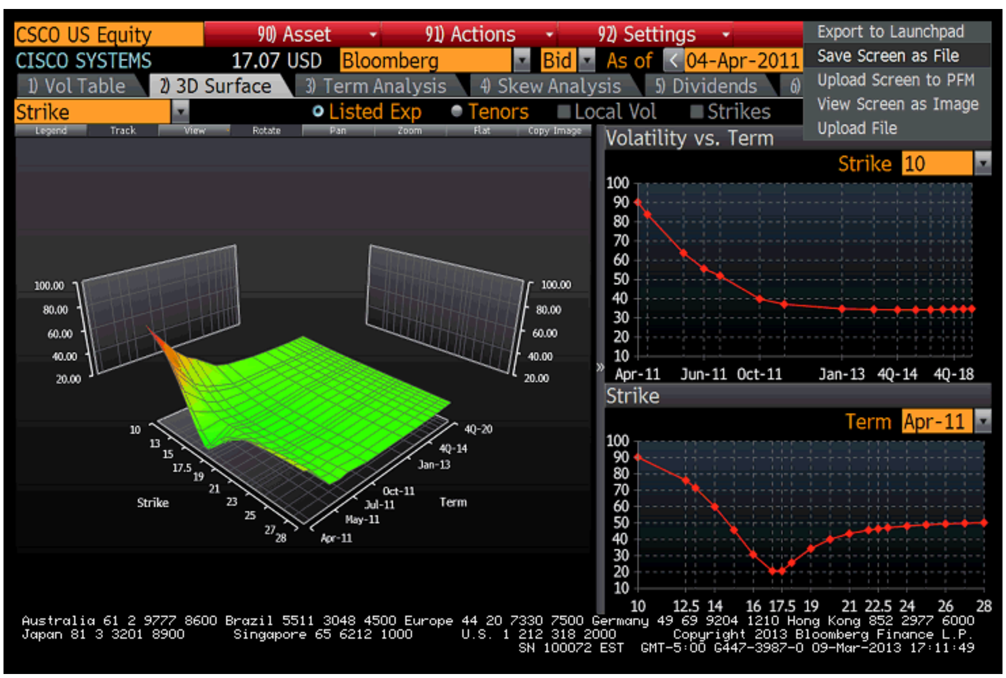
Strike Price	**Implied Volatility**	**Simulated Price**	**BSM Price**	**Abs(Sim - BSM)**
16.87	14.65%	0.168513	0.161877	0.006635
16.88	13.53%	0.151661	0.145026	0.006635
16.90	12.39%	0.134810	0.128175	0.006635
16.92	11.23%	0.117959	0.111325	0.006634
16.94	10.04%	0.101108	0.094485	0.006623
16.95	8.81%	0.084256	0.077695	0.006562
16.97	7.53%	0.067405	0.061103	0.006302
16.99	6.18%	0.050554	0.045117	0.005437
17.00	4.75%	0.033755	0.030531	0.003224
17.02	4.24%	0.020743	0.018414	0.002330
17.04	4.86%	0.014183	0.009619	0.004563
17.05	5.26%	0.009269	0.004240	0.005029
17.07	5.48%	0.005617	0.001542	0.004075
17.09	5.60%	0.003147	0.000455	0.002693
17.10	5.63%	0.001565	0.000107	0.001458
17.12	5.68%	0.000741	0.000020	0.000721
17.14	5.77%	0.000346	0.000003	0.000343
17.15	5.75%	0.000131	0.000000	0.000131
17.17	5.75%	0.000047	0.000000	0.000047
17.19	5.68%	0.000013	0.000000	0.000013

Table 5.6: Simulation summary of European call options on CSCO ($\sigma_{\mathrm{BSM}} = 3.56\%$).

We notice that our simulated ATM value of \$17.02 is very closed to the actual ATM value of \$17.07 from Bloomberg. Once again, our simulation result in Figure 5.6a shows a volatility smile as observed in real markets in Figure 5.6b, which confirms the plausibility of our model.



(a) Simulated implied volatility as a function of the strike price,



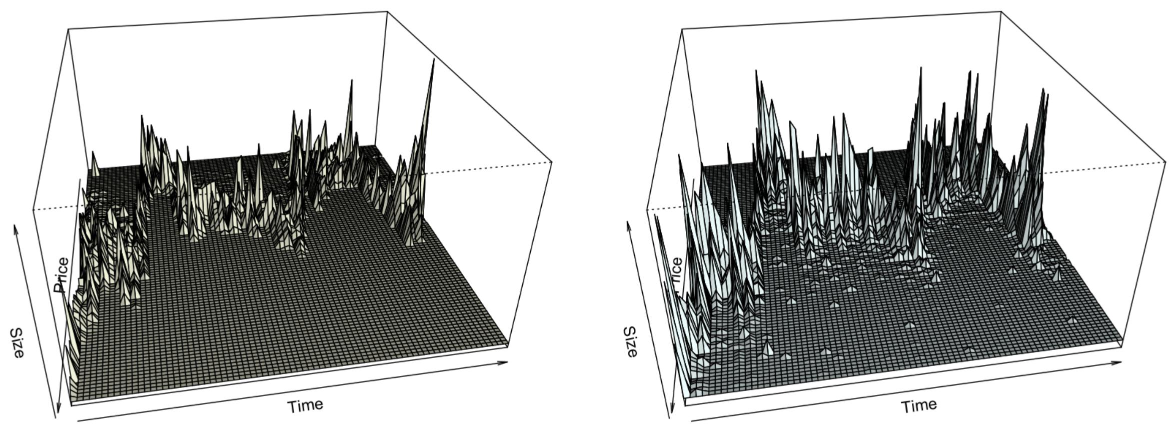
(b) Actual implied volatility as a function of the strike price from Bloomberg.

Figure 5.6: Implied volatility as a function of the strike price of CSCO.

5.1.3 Freeport-McMoRan Copper & Gold Inc. (FCX)

Let $p_{\text{FCX}_{\text{min}}}$ be the minimum price of USD 55.30 and $p_{\text{FCX}_{\text{max}}}$ be the maximum price of USD 56.27 during the core trading section from 9:30 a.m. to 4:00 p.m. ET on April 4, 2011. We partition the price range into one-cent bins and the time span into one-minute bins

$$p_{\text{FCX}_{\text{min}}} = \$55.30 = p_{-K} < p_{-K+1} < \cdots < p_{K-1} < p_K = p_{\text{FCX}_{\text{max}}} = \$56.27,$$
$$p_{K+1} - p_K = \Delta p \text{ (one cent)},$$
$$t_{i+1} - t_i = \Delta t \text{ (one minute)}.$$



(a) Buy Limit Orders

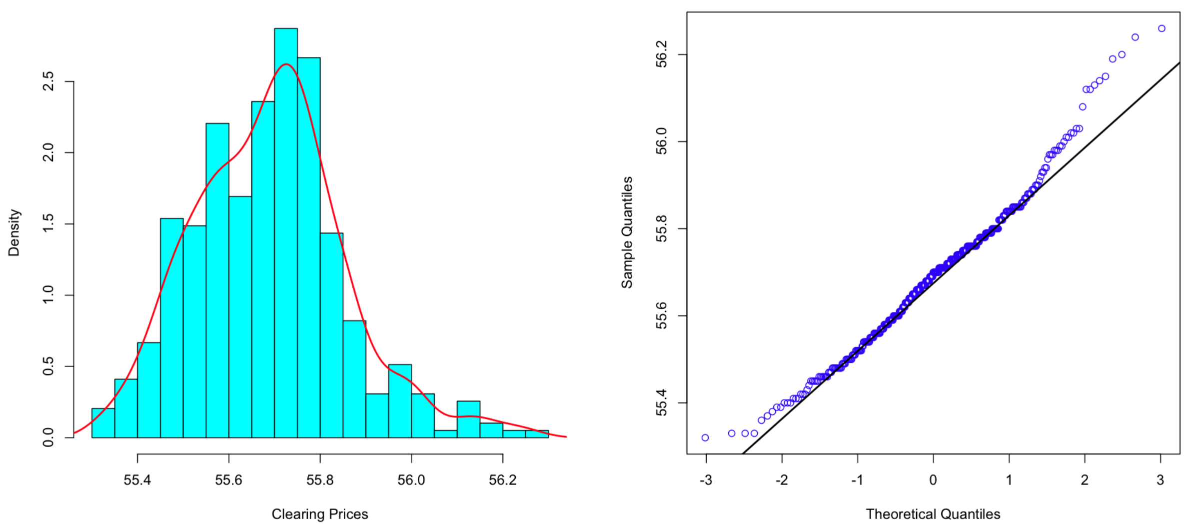
(b) Sell Limit Orders

Figure 5.7: The dynamics of Freeport-McMoRan Copper & Gold Inc.'s LOB on April 4, 2011.

By dividing the core trading section from 9:30 a.m. to 4:00 p.m. ET into one-minute intervals, we obtain 390 time intervals and present the summary statistics of $\pi_{\text{FCX}}(t)$ on April 4, 2011 in Table 5.7 as follow

Summary Statistics			
nobs	390	Sum	21718.87
NAs	0	SE Mean	0.0084
Minimum	55.30	LCL Mean	55.6729
Maximum	56.27	UCL Mean	55.7059
1. Quartile	55.57	Variance	0.0274
3. Quartile	55.78	Stdev	0.1655
Mean	55.6894	Skewness	0.4518
Median	55.7000	Kurtosis	0.5149

Table 5.7: Summary statistics of $\pi_{\text{FCX}}(t)$ (1 minute) on April 4, 2011.



(a) The histogram of $\pi_{\text{FCX}}(t)$ (b) Normal Q-Q plot of $\pi_{\text{FCX}}(t)$

Figure 5.8: $\pi_{\text{FCX}}(t)$ (1 minute) on April 4, 2011.

Similarly, we conduct the parameter estimation of our model based on the simulation steps over the limit order data of Freeport-McMoRan Copper & Gold Inc. (FCX) described in Chapter 4. From the observed data, the opening clearing price of FCX on April 4, 2011 was

$$\pi_{\text{FCX}}(0) = \$55.04.$$

Table 5.8: The parameter estimation of Freeport-McMoRan Copper & Gold Inc. (FCX).

Parameter	**Initial Value**	**Speed Mean Reversion**	**Volatility**	**Long-Term Value**
log Qtilde $p_{\min}$	21.342769	0.000419	0.110366	20.531192
log qtilde(-47 , t)	15.481197	0.000001	0.604595	14.247392
log qtilde(-46 , t)	15.042296	0.000030	0.612546	13.359617
log qtilde(-45 , t)	14.392255	0.000002	0.643072	13.079716
log qtilde(-44 , t)	15.182224	0.011336	0.637301	12.641851
log qtilde(-43 , t)	14.702251	0.000043	0.612104	13.509640
log qtilde(-42 , t)	14.964616	0.000000	0.618372	13.407195
log qtilde(-41 , t)	15.296233	0.000000	0.115532	14.375561
log qtilde(-40 , t)	14.634523	0.000000	0.128752	14.534841
log qtilde(-39 , t)	14.874816	0.000003	0.585809	13.562597
log qtilde(-38 , t)	14.971561	0.000000	0.624781	13.699307
log qtilde(-37 , t)	15.438612	0.000003	0.634346	14.648465
log qtilde(-36 , t)	15.404059	0.000450	0.070882	15.076001
log qtilde(-35 , t)	14.972677	0.006752	0.591331	11.716734
log qtilde(-34 , t)	15.338373	0.000691	0.082449	14.548242
log qtilde(-33 , t)	16.414784	0.000436	0.068979	15.785289
log qtilde(-32 , t)	15.762733	0.007588	0.625905	13.027832
log qtilde(-31 , t)	15.006262	0.007969	0.625152	12.615482
log qtilde(-30 , t)	16.062004	0.000672	0.046553	15.258914
log qtilde(-29 , t)	15.790570	0.000633	0.075738	14.401539
log qtilde(-28 , t)	15.832797	0.000626	0.060889	15.009158
log qtilde(-27 , t)	15.903307	0.005675	0.590436	12.315701
log qtilde(-26 , t)	16.863821	0.000517	0.086048	15.839305
log qtilde(-25 , t)	16.603124	0.000593	0.091820	15.672015
log qtilde(-24 , t)	16.049154	0.007762	0.639100	13.038188
log qtilde(-23 , t)	17.079603	0.000529	0.080888	15.079089
log qtilde(-22 , t)	16.812732	0.006647	0.646542	13.213829
log qtilde(-21 , t)	16.683348	0.000284	0.067110	15.663826
log qtilde(-20 , t)	16.782108	0.006602	0.638239	13.658835
log qtilde(-19 , t)	17.521722	0.000440	0.074204	15.987451
log qtilde(-18 , t)	17.382713	0.000423	0.095216	15.975168
log qtilde(-17 , t)	17.098560	0.001036	0.185294	14.766737

Continued on next page

Table 5.8 -- *Continued from previous page*

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log qtilde(-16 , t)	17.778879	0.000536	0.108052	16.950711
log qtilde(-15 , t)	17.567159	0.001291	0.172937	16.129606
log qtilde(-14 , t)	17.717686	0.000949	0.147489	16.341146
log qtilde(-13 , t)	17.428475	0.000000	0.235497	16.712824
log qtilde(-12 , t)	18.270389	0.000676	0.128987	16.884049
log qtilde(-11 , t)	18.010557	0.001235	0.168553	16.423634
log qtilde(-10 , t)	17.898276	0.000935	0.136291	15.733745
log qtilde(-9 , t)	17.869225	0.000561	0.130867	16.386437
log qtilde(-8 , t)	18.641931	0.000690	0.158519	16.923129
log qtilde(-7 , t)	18.328744	0.000893	0.174484	16.746471
log qtilde(-6 , t)	18.225163	0.000130	0.054689	17.170767
log qtilde(-5 , t)	18.802050	0.001047	0.153196	16.929257
log qtilde(-4 , t)	18.589870	0.000255	0.081105	17.517687
log qtilde(-3 , t)	18.489976	0.000355	0.099774	17.461956
log qtilde(-2 , t)	18.474253	0.000412	0.133048	17.533393
log qtilde(-1 , t)	19.092696	0.000263	0.083686	18.150120
log qtilde(0 , t)	18.591075	0.000298	0.088402	17.838259
log qtilde(1 , t)	18.668611	0.000244	0.068185	17.762427
log qtilde(2 , t)	19.167043	0.000330	0.104546	18.294185
log qtilde(3 , t)	18.641878	0.000326	0.124529	17.658630
log qtilde(4 , t)	18.629488	0.000219	0.069155	17.695462
log qtilde(5 , t)	18.646840	0.000198	0.059387	17.108353
log qtilde(6 , t)	18.987287	0.000582	0.147530	17.951744
log qtilde(7 , t)	18.323807	0.000199	0.061881	17.291157
log qtilde(8 , t)	18.480402	0.004169	0.661667	16.244294
log qtilde(9 , t)	18.844405	0.000654	0.169119	17.974880
log qtilde(10 , t)	18.121335	0.000602	0.123963	17.330255
log qtilde(11 , t)	18.092395	0.003884	0.683890	17.308709
log qtilde(12 , t)	18.221349	0.006893	0.714487	14.370714
log qtilde(13 , t)	18.509541	0.004471	0.650300	14.227250
log qtilde(14 , t)	17.594201	0.005502	0.658283	13.704121
log qtilde(15 , t)	17.696825	0.005240	0.651581	13.965227
log qtilde(16 , t)	17.676497	0.000004	0.590934	16.928750

Continued on next page

Table 5.8 -- *Continued from previous page*

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log qtilde(17 , t)	17.827842	0.004394	0.641898	14.129905
log qtilde(18 , t)	16.993706	0.003911	0.586657	13.484149
log qtilde(19 , t)	17.128569	0.000386	0.108688	16.455004
log qtilde(20 , t)	17.378848	0.000004	0.590173	16.696178
log qtilde(21 , t)	16.460827	0.005447	0.631111	12.170059
log qtilde(22 , t)	16.763343	0.006631	0.635934	12.950877
log qtilde(23 , t)	16.491581	0.006270	0.638216	13.281166
log qtilde(24 , t)	16.975362	0.005175	0.623461	14.483164
log qtilde(25 , t)	15.906867	0.006871	0.618511	13.473526
log qtilde(26 , t)	16.171248	0.004687	0.588178	12.795448
log qtilde(27 , t)	16.268707	0.005665	0.622917	13.141960
log qtilde(28 , t)	15.420697	0.000039	0.594922	15.174664
log qtilde(29 , t)	15.640161	0.005629	0.587480	12.154008
log qtilde(30 , t)	15.638398	0.005711	0.587541	12.387739
log qtilde(31 , t)	15.977815	0.010889	0.681097	13.111441
log qtilde(32 , t)	15.392021	0.090593	0.988603	14.154002
log qtilde(33 , t)	15.140581	0.004122	0.587500	14.712424
log qtilde(34 , t)	15.494267	0.007937	0.650559	12.882318
log qtilde(35 , t)	14.504327	0.005716	0.563686	11.468452
log qtilde(36 , t)	14.698096	0.121337	1.012161	13.870142
log qtilde(37 , t)	14.514844	0.005997	0.568579	11.982703
log qtilde(38 , t)	14.980826	0.001373	0.110337	14.194388
log qtilde(39 , t)	11.790966	0.005892	0.557715	10.219186
log qtilde(40 , t)	13.717194	0.007019	0.567217	11.506779
log qtilde(41 , t)	12.540944	0.001641	0.084405	12.991555
log qtilde(42 , t)	13.870781	0.025487	0.665795	12.494107
log qtilde(43 , t)	10.587392	0.006146	0.578922	8.952042
log qtilde(44 , t)	14.058075	0.000000	0.594852	12.368589
log qtilde(45 , t)	14.104492	0.000000	0.595345	12.436592
log qtilde(46 , t)	13.369445	0.009383	0.594960	9.925506
log qtilde(47 , t)	14.671004	0.000003	0.575853	13.200918
log qtilde(48 , t)	14.285531	0.000007	0.552645	12.774948
Clearing Price $\pi(0)$	55.04			

We simulate the lognormal models and the clearing price process $\pi(t)$ based on the formulas (4.8), (4.9) and (4.10) respectively, and then we approximate the value of a European call option with 500 scenarios

$$C(\psi) = \frac{1}{N} \sum_{\omega=1}^{N} \text{max}(\pi(T, \omega) - \psi, 0),$$

$$\begin{aligned} \text{where :} \quad & N = 500, \\ & \psi = \text{a strike price of the option}, \\ & T = \frac{1}{365}, \text{ that is equal to one day.} \end{aligned}$$

Likewise, we compare our simulated option price with the approximated value of a European call option for the same underlying stock from the Black-Scholes-Merton model with a constant σ_{BSM} of 38.21%. As a result, we find that our simulated prices are closed to the BSM values in Table 5.9, which proves the accuracy of our model.

Strike Price	Implied Volatility	Simulated Price	BSM Price	Abs(Sim - BSM)
55.83	29.59%	0.083264	0.155351	0.072087
55.88	28.95%	0.069347	0.142651	0.073304
55.94	28.40%	0.057520	0.130775	0.073255
56.00	27.87%	0.047256	0.119691	0.072435
56.05	27.33%	0.038270	0.109365	0.071095
56.11	26.82%	0.030646	0.099763	0.069117
56.16	26.34%	0.024261	0.090851	0.066590
56.22	25.98%	0.019341	0.082596	0.063256
56.28	25.55%	0.015015	0.074964	0.059949
56.33	25.09%	0.011380	0.067921	0.056541
56.39	24.73%	0.008676	0.061434	0.052758
56.44	24.40%	0.006539	0.055470	0.048932
56.50	24.04%	0.004827	0.049999	0.045172
56.55	23.75%	0.003564	0.044989	0.041426
56.61	23.54%	0.002648	0.040411	0.037763
56.67	23.30%	0.001927	0.036234	0.034307
56.72	23.26%	0.001487	0.032432	0.030946
56.78	23.47%	0.001254	0.028978	0.027724
56.83	23.61%	0.001031	0.025845	0.024814
56.89	23.63%	0.000808	0.023011	0.022203

Table 5.9: Simulation summary of European call options on FCX ($\sigma_{\text{BSM}} = 38.21\%$).

Comparing to other favorable results, we notice that our simulated ATM value of \$56.72 on FCX is not quite closed to the actual ATM value of \$55.78 from Bloomberg, but its volatility of 23.26% is reasonable. With that said, once again our simulation result in Figure 5.9a shows a volatility smile as observed in real markets in Figure 5.9b, which confirms the plausibility of our model.

(a) Simulated implied volatility as a function of the strike price,

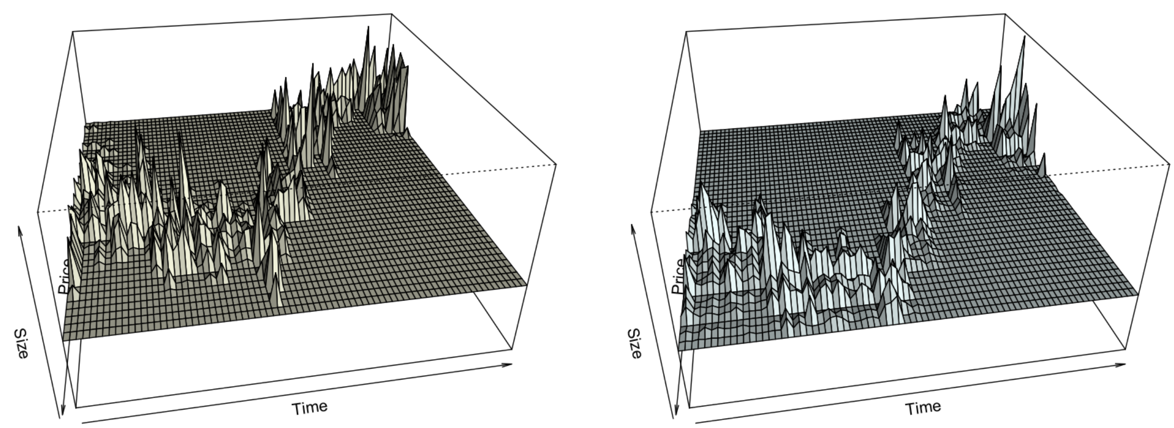
(b) Actual implied volatility as a function of the strike price from Bloomberg.

Figure 5.9: Implied volatility as a function of the strike price of FCX.

5.1.4 Oracle Corporation (ORCL)

Let $p_{\text{ORCL}_{\text{min}}}$ be the minimum price of USD 34.01 and $p_{\text{ORCL}_{\text{max}}}$ be the maximum price of USD 34.43 during the core trading section from 9:30 a.m. to 4:00 p.m. ET on April 4, 2011. We partition the price range into one-cent bins and the time span into one-minute bins

$$p_{\text{ORCL}_{\text{min}}} = \$34.01 = p_{-K} < p_{-K+1} < \cdots < p_{K-1} < p_K = p_{\text{ORCL}_{\text{max}}} = \$34.43,$$
$$p_{K+1} - p_K = \Delta p \text{ (one cent)},$$
$$t_{i+1} - t_i = \Delta t \text{ (one minute)}.$$



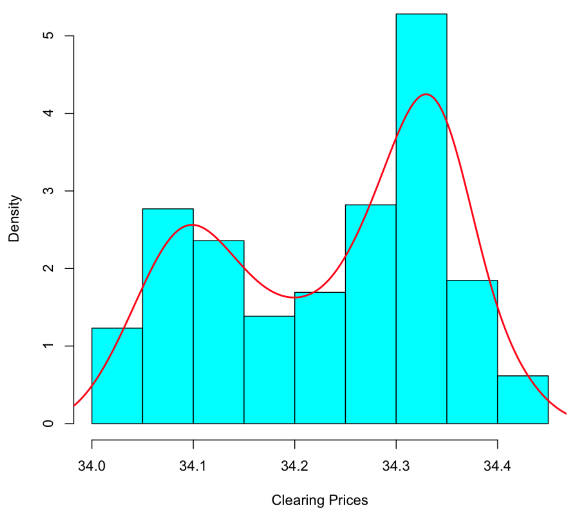
(a) Buy Limit Orders (b) Sell Limit Orders

Figure 5.10: The dynamics of Oracle Corporation's LOB on April 4, 2011.

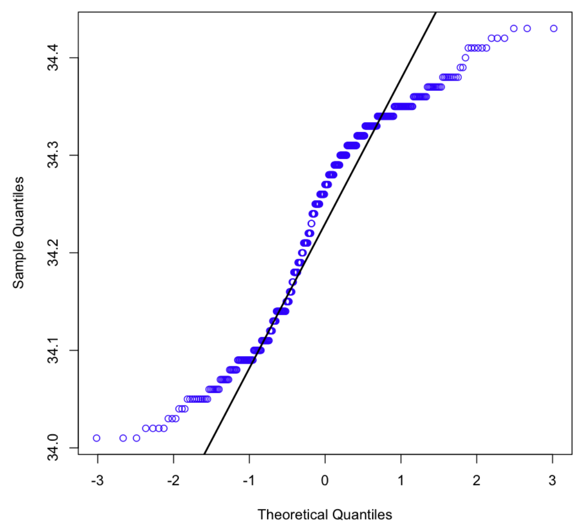
By dividing the core trading section from 9:30 a.m. to 4:00 p.m. ET into one-minute intervals, we obtain 390 time intervals and present the summary statistics of $\pi_{\text{ORCL}}(t)$ on April 4, 2011 in Table 5.10 as follow

Summary Statistics			
nobs	390	Sum	13351.88
NAs	0	SE Mean	0.0057
Minimum	34.01	LCL Mean	34.2243
Maximum	34.43	UCL Mean	34.2469
1. Quartile	34.13	Variance	0.0129
3. Quartile	34.33	Stdev	0.1134
Mean	34.2356	Skewness	-0.3230
Median	34.2700	Kurtosis	-1.2561

Table 5.10: Summary statistics of $\pi_{\text{ORCL}}(t)$ (1 minute) on April 4, 2011.



(a) The histogram of $\pi_{\text{ORCL}}(t)$



(b) Normal Q-Q plot of $\pi_{\text{ORCL}}(t)$

Figure 5.11: $\pi_{\text{ORCL}}(t)$ (1 minute) on April 4, 2011.

Similarly, we conduct the parameter estimation of our model based on the simulation steps over the limit order data of Oracle Corporation (ORCL) described in Chapter 4. From the observed data, the opening clearing price of ORCL on April 4, 2011 was

$$\pi_{\text{ORCL}}(0) = \$34.01.$$

Table 5.11: The parameter estimation of Oracle Corporation (ORCL).

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log Qtilde $p_{\min}$	20.817943	0.000314	0.082737	18.447554
log qtilde(-20 , t)	14.964775	0.003341	0.061809	12.950032
log qtilde(-19 , t)	14.128055	0.005946	0.550274	10.224772
log qtilde(-18 , t)	13.912699	0.004149	0.514837	9.372386
log qtilde(-17 , t)	13.414800	0.004569	0.524047	11.748233
log qtilde(-16 , t)	14.889440	0.003911	0.513691	11.359074
log qtilde(-15 , t)	14.702663	0.003619	0.512694	10.599464
log qtilde(-14 , t)	14.786092	0.003592	0.516965	11.542149
log qtilde(-13 , t)	15.811493	0.003465	0.512261	11.827624
log qtilde(-12 , t)	15.905748	0.003069	0.514886	12.220714
log qtilde(-11 , t)	16.071278	0.000035	0.510396	15.297579
log qtilde(-10 , t)	16.126856	0.000013	0.508959	15.547992
log qtilde(-9 , t)	17.031239	0.000000	0.610628	15.475524
log qtilde(-8 , t)	17.170371	0.000265	0.092721	15.936148
log qtilde(-7 , t)	17.104169	0.000000	0.505093	16.389672
log qtilde(-6 , t)	17.486061	0.000421	0.119075	15.867232
log qtilde(-5 , t)	18.200920	0.000000	0.571880	16.917104
log qtilde(-4 , t)	18.090539	0.000758	0.139876	14.246055
log qtilde(-3 , t)	18.140405	0.000381	0.117286	16.093006
log qtilde(-2 , t)	18.609524	0.000411	0.096216	14.176791
log qtilde(-1 , t)	18.559878	0.000346	0.101906	16.239163
log qtilde(0 , t)	18.448956	0.000146	0.063540	16.668114
log qtilde(1 , t)	18.433650	0.000957	0.132817	16.449237
log qtilde(2 , t)	19.122119	0.000317	0.091608	17.068642
log qtilde(3 , t)	18.590020	0.000537	0.130293	16.786755
log qtilde(4 , t)	18.766931	0.000282	0.073521	16.654372
log qtilde(5 , t)	19.204290	0.000291	0.087546	16.321255
log qtilde(6 , t)	18.472854	0.000275	0.076941	15.937492
log qtilde(7 , t)	18.569892	0.000785	0.166135	15.868562
log qtilde(8 , t)	18.582991	0.000254	0.052776	15.679859
log qtilde(9 , t)	18.602124	0.000296	0.083468	16.129950
log qtilde(10 , t)	17.836014	0.000264	0.076907	14.581827

Continued on next page

Table 5.11 -- *Continued from previous page*

Parameter	**Initial Value**	**Speed Mean Reversion**	**Volatility**	**Long-Term Value**
log qtilde(11 , t)	18.028523	0.000017	0.369180	15.452644
log qtilde(12 , t)	18.182111	0.000466	0.092100	14.818413
log qtilde(13 , t)	17.297709	0.000258	0.061787	15.096508
log qtilde(14 , t)	17.205374	0.000453	0.061377	14.102356
log qtilde(15 , t)	17.103052	0.003592	0.494076	11.090121
log qtilde(16 , t)	17.154217	0.000335	0.074145	14.231777
log qtilde(17 , t)	16.142578	0.007116	0.555297	11.435039
log qtilde(18 , t)	16.260833	0.004599	0.497537	10.885186
log qtilde(19 , t)	15.659319	0.005431	0.508565	10.831284
log qtilde(20 , t)	15.901217	0.336116	0.774668	13.106000
log qtilde(21 , t)	15.499524	0.000298	0.990652	11.283025
Clearing Price $\pi(0)$	34.01			

We simulate the lognormal models and the clearing price process $\pi(t)$ based on the formulas (4.8), (4.9) and (4.10) respectively, and then we approximate the value of a European call option with 500 scenarios

$$C(\psi) = \frac{1}{N}\sum_{\omega=1}^{N} \text{max}(\pi(T,\omega) - \psi, 0),$$

$$\begin{aligned} \text{where :} \quad & N = 500, \\ & \psi = \text{a strike price of the option}, \\ & T = \frac{1}{365}, \text{ that is equal to one day.} \end{aligned}$$

Likewise, we compare our simulated option price with the approximated value of a European call option for the same underlying stock from the Black-Scholes-Merton model with a constant σ_{BSM} of 13.75%. As a result, we find that our simulated prices are very closed to the BSM values in Table 5.12, which once again proves the accuracy of our model.

Strike Price	Implied Volatility	Simulated Price	BSM Price	Abs(Sim - BSM)
33.77	23.55%	0.303660	0.249569	0.054090
33.81	21.90%	0.269986	0.222462	0.047524
33.84	20.21%	0.236313	0.196725	0.039589
33.87	18.44%	0.202641	0.172498	0.030143
33.91	16.61%	0.168997	0.149906	0.019090
33.94	14.70%	0.135593	0.129049	0.006544
33.98	12.70%	0.102564	0.109999	0.007435
34.01	10.58%	0.070289	0.092794	0.022505
34.04	10.36%	0.053797	0.077440	0.023642
34.08	11.16%	0.046576	0.063905	0.017329
34.11	11.90%	0.040417	0.052127	0.011710
34.14	12.56%	0.035009	0.042013	0.007004
34.18	13.10%	0.029918	0.033446	0.003528
34.21	13.53%	0.025220	0.026291	0.001071
34.25	13.84%	0.020764	0.020400	0.000364
34.28	14.06%	0.016782	0.015620	0.001162
34.31	14.19%	0.013205	0.011800	0.001405
34.35	14.23%	0.010103	0.008792	0.001311
34.38	14.14%	0.007314	0.006459	0.000855
34.41	13.86%	0.004869	0.004679	0.000190

Table 5.12: Simulation summary of European call options on ORCL ($\sigma_{\text{BSM}} = 13.75\%$).

We also notice that our simulated ATM value of $34.04 is very closed to the actual ATM value of $34.10 from Bloomberg. Most importantly, our simulation result in Figure 5.12a shows a volatility smile as observed in real markets in Figure 5.12b, which confirms the plausibility of our model.

(a) Simulated implied volatility as a function of the strike price,

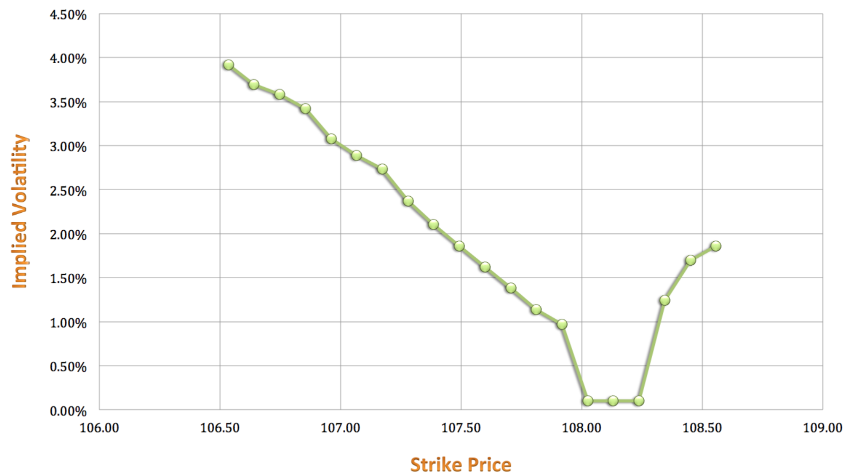
(b) Actual implied volatility as a function of the strike price from Bloomberg.

Figure 5.12: Implied volatility as a function of the strike price of ORCL.

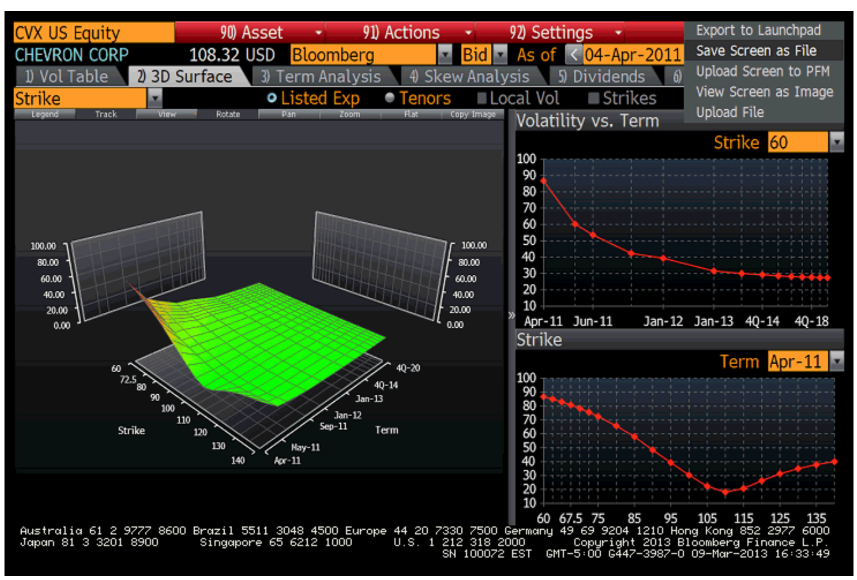
5.2 Open Issues for Extensive Studies

5.2.1 Chevron Corporation (CVX)

Under the similar simulation scheme, we obtain the parameter estimation and simulation result of European call options on Chevron Corporation (CVX) in Appendix C.



(a) Simulated implied volatility as a function of the strike price,



(b) Actual implied volatility as a function of the strike price from Bloomberg.

Figure 5.13: Implied volatility as a function of the strike price of CVX.

5.2.2 JPMorgan Chase & Co. (JPM)

We present the parameter estimation and simulation result of European call options on JP-Morgan Chase & Co. (JPM) in Appendix D.

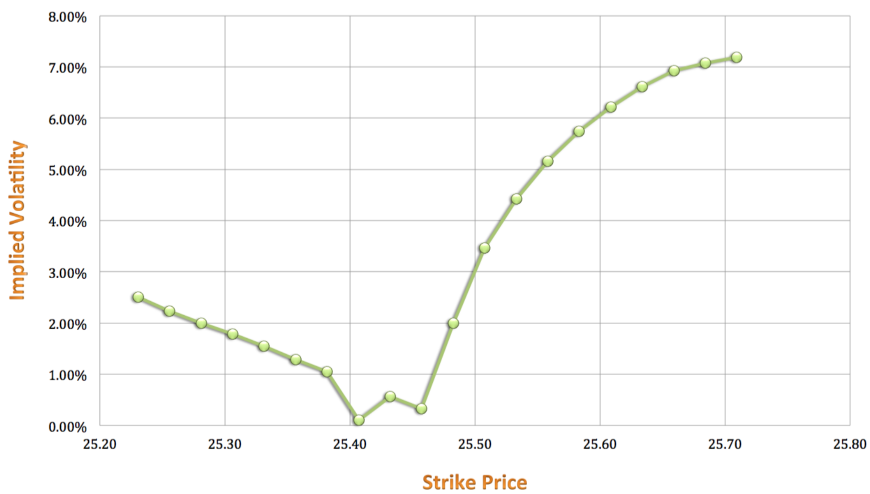
(a) Simulated implied volatility as a function of the strike price,

(b) Actual implied volatility as a function of the strike price from Bloomberg.

Figure 5.14: Implied volatility as a function of the strike price of JPM.

5.2.3 Microsoft Corporation (MSFT)

We present the parameter estimation and simulation result of European call options on Microsoft Corporation (MSFT) in Appendix E.



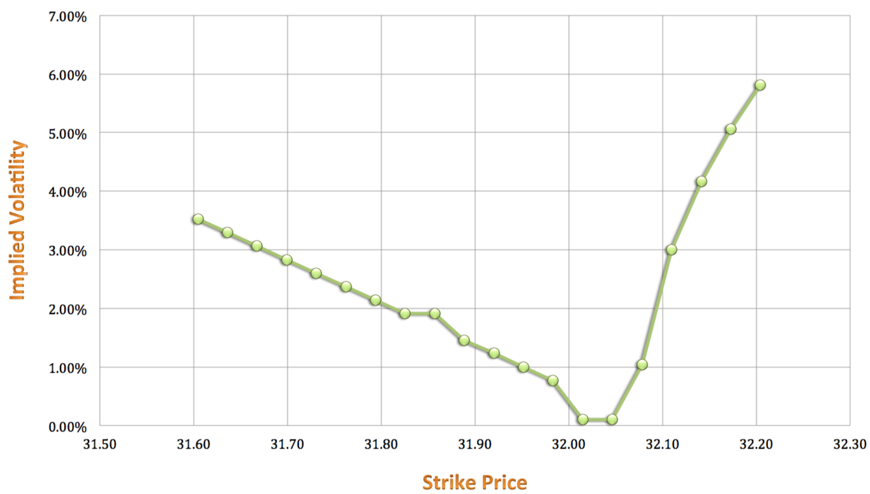
(a) Simulated implied volatility as a function of the strike price,

(b) Actual implied volatility as a function of the strike price from Bloomberg.

Figure 5.15: Implied volatility as a function of the strike price of MSFT.

5.2.4 Wells Fargo & Co. (WFC)

We present the parameter estimation and simulation result of European call options on Wells Fargo & Co. (WFC) in Appendix F.



(a) Simulated implied volatility as a function of the strike price,

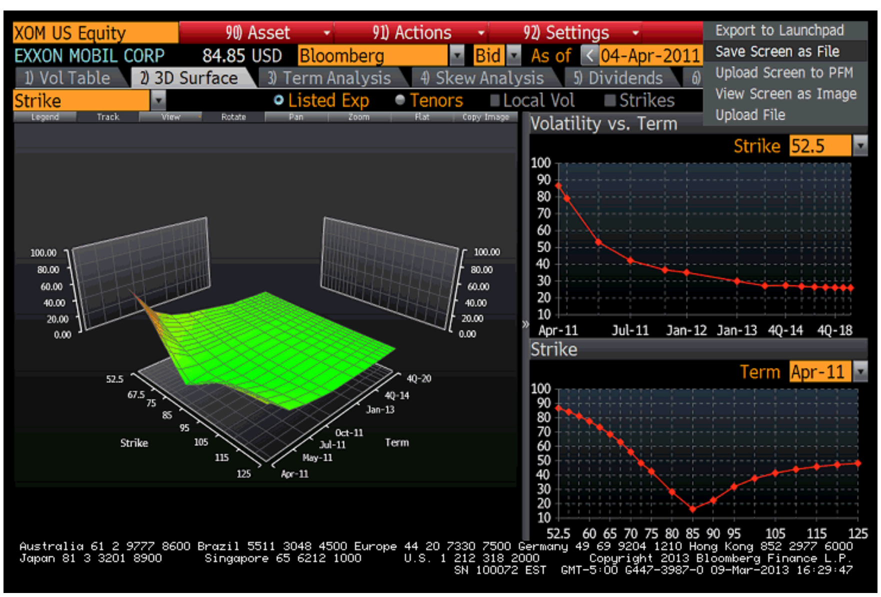
(b) Actual implied volatility as a function of the strike price from Bloomberg.

Figure 5.16: Implied volatility as a function of the strike price of WFC.

5.2.5 Exxon Mobil Corporation (XOM)

We present the parameter estimation and simulation result of European call options on Exxon Mobil Corporation (XOM) in Appendix G.

(a) Simulated implied volatility as a function of the strike price,



(b) Actual implied volatility as a function of the strike price from Bloomberg.

Figure 5.17: Implied volatility as a function of the strike price of XOM.

In general, although these simulated results are less favorable as expected, we observe that the simulated implied volatilities from our model did change for each strike price as observed in real markets. Most importantly, our simulated implied volatility as a function of the strike price are still in the U-shape curve while our simulated ATM values are also closed to the actual ATM values from Bloomberg.

Firm	**Ticker**	**Simulated Value**	**Actual Value**
Chevron Corporation	CVX	108.13	108.32
JPMorgan Chase & Co.	JPM	46.18	46.28
Microsoft Corporation	MSFT	25.41	25.56
Wells Fargo & Co.	WFC	32.05	31.79
Exxon Mobil Corporation	XOM	84.46	84.85

Table 5.13: Comparison between simulated and actual at-the-money (ATM) values.

From both favorable and open-issue simulation results, we strongly recommend extensive studies to investigate further. In the next chapter, we will present our conclusions and suggest some extension to our research.

Chapter 6

Conclusions

In our research, we present an empirical study of the no-arbitrage liquidity model in financial markets. In our model, the limit order flow completely determines the clearing prices of the assets. By employing a Brownian sheet and the Ito-Wentzell formula, we model both the net demand curve and the clearing price process under the risk-neutral measure. We develop a semi-relative no-arbitrage liquidity model focusing on the absolute quantities at the limit price p when the clearing price is $\pi(t)$ with a purpose to keep the clearing prices bounded when we simulate the prices. We then simulate our model with HF aggregate limit order data provided by NYSE Arcabook, including nine stocks from four different industries, such as energy (CVX, XOM), financial banking (JPM, WFC), materials and mining (ABX, FCX) and technology (CSCO, MSFT, ORCL). Under our simulation methodology, we approximate the option price and calculate the implied volatility as a function of the strike price for each European option. We also compare our simulated European option prices to the values from the Black-Scholes-Merton model as well as compare our simulated volatility smiles to the actual ones from Bloomberg.

In general, we accomplished several objectives of our research. Firstly, we analyzed a massive data set of 300GB with more than 5 billion data rows. We also successfully developed and automated the parallelization algorithm to handle this data set without using any super computers. Our parallelization methodology will be very useful for researchers who are planning to work on a massive data. Secondly, we specified a very complicated model with a positive demand density. Lastly and most importantly, our model generated volatility smiles as observed in real markets from Bloomberg, which the Black-Scholes model does not.

In conclusion, we observe that both favorable and open-issue results objectively confirm the potentiality and plausibility of our model. Therefore, we highly suggest that extensive studies continue to be investigated further. We recommend these following extensions to our research. Firstly, we have not considered the effect of the trading date on our model. For example, Chordia et al. (2001) conclude that the market liquidity decreases and the trading activity slows on Friday while the opposite situation occurs on Tuesday. We can apply our model on the data of different trading date to investigate the behavior of our model under the change of the market liquidity. Secondly, in order to expand the price range between $p_{\max}$ and $p_{\min}$, we can extend our estimation period to a longer time span, i.e. one week, but the modifications for Δt and Δp will be needed. Lastly, developing a different empirical model based on the theory in German and Schellhorn (2012) is strongly encouraged.

Appendix A

Cancelation Methodology for Limit Orders

In this section, we show that under our methodology, the cancelation rates are approximately up to 20%, consistent with Gould et al. (2012).

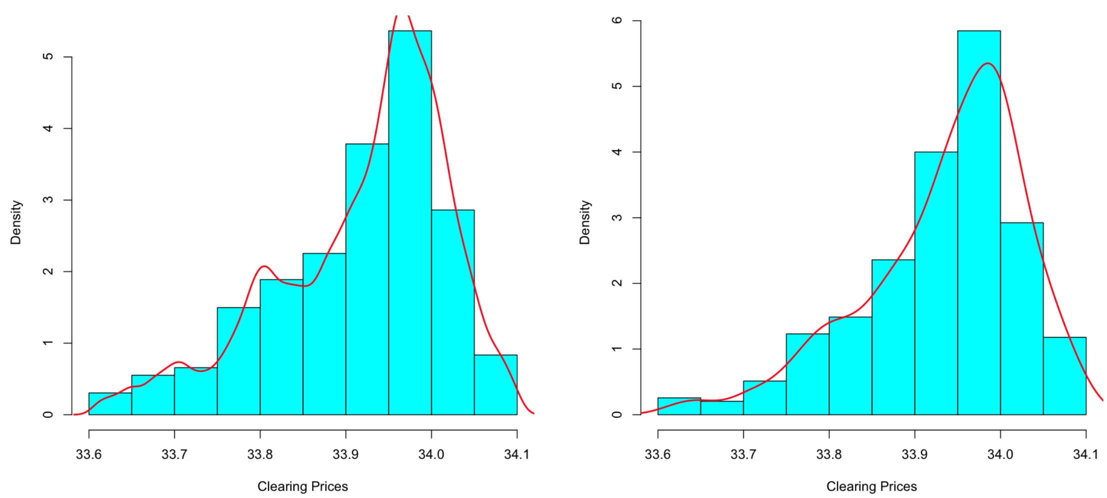
Firm	Ticker	Buy Limit Orders	Sell Limit Orders
Barrick Gold Corporation	ABX	6.52%	13.22%
Cisco Systems, Inc.	CSCO	14.22%	10.78%
Chevron Corporation	CVX	16.19%	21.60%
Freeport-McMoRan Copper & Gold Inc.	FCX	5.63%	5.11%
JPMorgan Chase & Co.	JPM	3.39%	3.10%
Microsoft Corporation	MSFT	3.27%	3.63%
Oracle Corporation	ORCL	8.87%	4.37%
Wells Fargo & Company	WFC	1.39%	1.39%
Exxon Mobil Corporation	XOM	12.03%	4.11%

Table A.1: Cancelation Percentages on April 4, 2011.

Appendix B

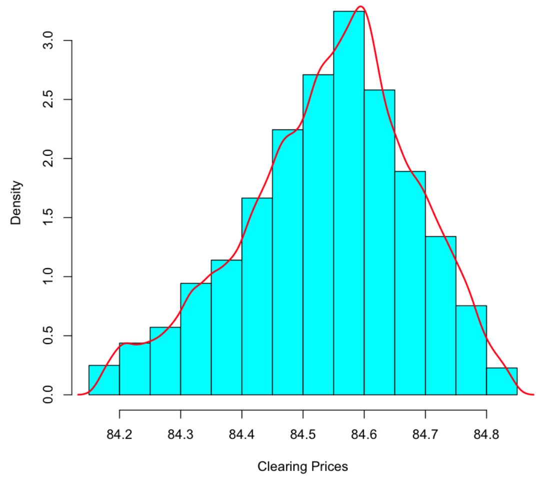
Selection Methodology for Clearing Prices

In this section, we show that the selection methodology for clearing prices performs properly by comparing the density of all clearing prices during a trading day to the density of 390 selected clearing prices in one-minute intervals.

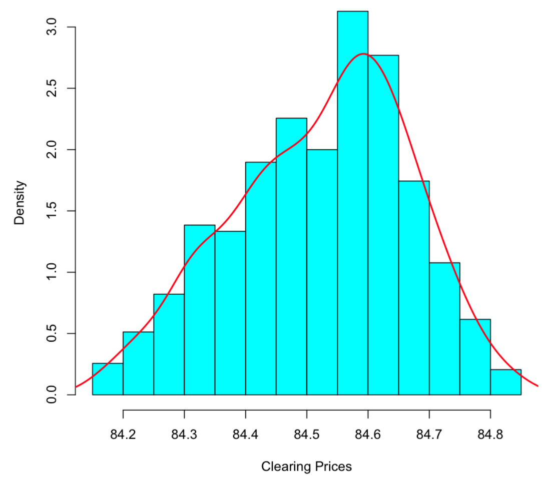


(a) All $\pi_{\text{ORCL}}(t)$ on April 1, 2011 (b) Selected $\pi_{\text{ORCL}}(t)$ (1 minute)

Figure B.1: Clearing Prices of Oracle Corporation (ORCL) on April 1, 2011.



(a) All $\pi_{\text{XOM}}(t)$ on April 1, 2011



(b) Selected $\pi_{\text{XOM}}(t)$ (1 minute)

Figure B.2: Clearing Prices of Exxon Mobil Corporation (XOM) on April 1, 2011.

In both Figures B.1 and B.2, the density of 390 selected clearing prices $\pi(t)$ in one-minute intervals almost reflects the density of all clearing prices during a trading day on April 1, 2011, which confirms the accuracy of our selection methodology.

Appendix C

Chevron Corporation (CVX)

Table C.1: The parameter estimation of Chevron Corporation (CVX).

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log Qtilde $p_{\min}$	20.733738	0.000431	0.113718	20.530805
log qtilde(-19 , t)	17.182796	0.008183	0.679294	13.703835
log qtilde(-18 , t)	16.444898	0.007559	0.642463	12.983362
log qtilde(-17 , t)	17.423108	0.005186	0.645837	14.527293
log qtilde(-16 , t)	17.362291	0.005919	0.645048	13.456756
log qtilde(-15 , t)	16.857264	0.008150	0.667810	13.586143
log qtilde(-14 , t)	17.690817	0.005932	0.669535	13.607661
log qtilde(-13 , t)	17.562086	0.006892	0.677239	13.981679
log qtilde(-12 , t)	17.306328	0.007926	0.686636	13.888791
log qtilde(-11 , t)	17.814396	0.004860	0.647446	13.787795
log qtilde(-10 , t)	17.769437	0.004818	0.644774	13.840713
log qtilde(-9 , t)	17.592469	0.004794	0.634291	13.730427
log qtilde(-8 , t)	18.120373	0.000544	0.150232	17.317198
log qtilde(-7 , t)	18.093579	0.000538	0.135614	17.283407
log qtilde(-6 , t)	17.954902	0.000833	0.162096	16.618286
log qtilde(-5 , t)	18.269227	0.004352	0.652085	14.539546
log qtilde(-4 , t)	17.617329	0.001049	0.162016	15.682886
log qtilde(-3 , t)	18.408764	0.004048	0.652449	14.699822
log qtilde(-2 , t)	18.278773	0.000612	0.168386	17.795570

Continued on next page

Table C.1 -- *Continued from previous page*

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log qtilde(-1 , t)	17.779247	0.004973	0.659605	14.222139
log qtilde(0 , t)	18.383714	0.000462	0.128214	17.947777
log qtilde(1 , t)	18.449070	0.004822	0.673290	14.441735
log qtilde(2 , t)	17.832327	0.001217	0.181675	15.900499
log qtilde(3 , t)	18.338877	0.000523	0.159714	17.858357
log qtilde(4 , t)	18.434900	0.000535	0.147512	17.949424
log qtilde(5 , t)	17.595945	0.001290	0.127987	15.756525
log qtilde(6 , t)	18.206749	0.000444	0.111749	18.079424
log qtilde(7 , t)	17.812167	0.000568	0.112426	17.687691
log qtilde(8 , t)	18.127185	0.000412	0.103470	17.977597
log qtilde(9 , t)	18.112632	0.000409	0.097995	17.894223
log qtilde(10 , t)	17.598446	0.000760	0.122457	16.976657
log qtilde(11 , t)	17.744695	0.000750	0.181342	17.844755
log qtilde(12 , t)	17.832732	0.000575	0.151716	17.974473
log qtilde(13 , t)	17.157868	0.005359	0.671496	14.195948
log qtilde(14 , t)	17.318928	0.000474	0.123029	17.617184
log qtilde(15 , t)	17.328578	0.000436	0.098223	17.843150
log qtilde(16 , t)	16.716594	0.001716	0.227349	15.158334
log qtilde(17 , t)	17.134242	0.000468	0.111041	17.312948
log qtilde(18 , t)	17.341240	0.000018	0.179893	17.760250
log qtilde(19 , t)	16.090711	0.000862	0.134524	14.229198
log qtilde(20 , t)	16.845882	0.000909	0.120421	17.016226
Clearing Price $\pi(0)$	108.32			

Strike Price	Implied Volatility	Simulated Price	BSM Price	Abs(Sim - BSM)
106.54	3.92%	1.596434	1.815980	0.219546
106.64	3.69%	1.490005	1.716880	0.226875
106.75	3.58%	1.383576	1.619173	0.235597
106.85	3.42%	1.277147	1.523049	0.245902
106.96	3.08%	1.170718	1.428706	0.257988
107.07	2.89%	1.064289	1.336353	0.272064
107.17	2.73%	0.957860	1.246201	0.288341
107.28	2.37%	0.851431	1.158466	0.307034
107.39	2.10%	0.745002	1.073359	0.328357
107.49	1.86%	0.638954	0.991090	0.352135
107.60	1.62%	0.533235	0.911855	0.378620
107.71	1.38%	0.427516	0.835841	0.408325
107.81	1.14%	0.322116	0.763213	0.441097
107.92	0.97%	0.221997	0.694120	0.472123
108.03	0.10%	0.143213	0.628681	0.485469
108.13	0.10%	0.087694	0.566994	0.479299
108.24	0.10%	0.046922	0.509122	0.462200
108.34	1.24%	0.017453	0.455101	0.437647
108.45	1.70%	0.003912	0.404933	0.401021
108.56	1.86%	0.000439	0.358589	0.358150

Table C.2: Simulation summary of European call options on CVX ($\sigma_{\mathrm{BSM}} = 20.66\%$).

Appendix D

JPMorgan Chase & Co. (JPM)

Table D.1: The parameter estimation of JPMorgan Chase & Co. (JPM).

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log Qtilde $p_{\min}$	21.207143	0.000335	0.067913	19.375692
log qtilde(-22 , t)	15.148200	0.012265	0.692026	12.573735
log qtilde(-21 , t)	14.319269	0.000001	0.516227	14.142388
log qtilde(-20 , t)	14.488253	0.000837	0.089761	14.726342
log qtilde(-19 , t)	14.020625	0.002879	0.563182	15.193561
log qtilde(-18 , t)	13.896711	0.004845	0.541899	10.526586
log qtilde(-17 , t)	15.111649	0.005109	0.574527	11.753339
log qtilde(-16 , t)	15.582302	0.006269	0.616654	12.566185
log qtilde(-15 , t)	15.518937	0.005906	0.613491	11.771751
log qtilde(-14 , t)	15.614080	0.003878	0.545537	12.160427
log qtilde(-13 , t)	15.643054	0.005945	0.587320	12.423775
log qtilde(-12 , t)	16.836169	0.004105	0.593474	13.117593
log qtilde(-11 , t)	16.874830	0.000696	0.049832	15.557006
log qtilde(-10 , t)	16.920530	0.005474	0.631115	14.313943
log qtilde(-9 , t)	17.045476	0.010682	0.707831	13.706754
log qtilde(-8 , t)	18.239737	0.003333	0.598471	14.937353
log qtilde(-7 , t)	18.287354	0.000806	0.129458	15.058942
log qtilde(-6 , t)	18.198619	0.000203	0.057114	16.286656
log qtilde(-5 , t)	18.932514	0.000216	0.061522	17.220356

Continued on next page

Table D.1 -- *Continued from previous page*

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log qtilde(-4 , t)	18.849365	0.000140	0.044758	17.030586
log qtilde(-3 , t)	18.986432	0.000173	0.048023	17.238880
log qtilde(-2 , t)	18.835315	0.000050	0.574913	17.265765
log qtilde(-1 , t)	19.729669	0.000171	0.051055	17.835292
log qtilde(0 , t)	19.354859	0.000178	0.055160	17.535929
log qtilde(1 , t)	19.421137	0.000997	0.141108	17.512428
log qtilde(2 , t)	19.859876	0.000798	0.132822	17.591067
log qtilde(3 , t)	19.058986	0.000551	0.091206	16.725487
log qtilde(4 , t)	19.068462	0.000676	0.094689	16.838664
log qtilde(5 , t)	19.068530	0.000446	0.073401	15.962279
log qtilde(6 , t)	19.183209	0.000482	0.089145	17.465233
log qtilde(7 , t)	18.197237	0.001939	0.161152	17.390998
log qtilde(8 , t)	18.288106	0.000231	0.040867	16.727005
log qtilde(9 , t)	18.482443	0.000338	0.048190	15.656034
log qtilde(10 , t)	17.309087	0.000274	0.040495	15.870505
log qtilde(11 , t)	17.352386	0.000872	0.076134	15.287708
log qtilde(12 , t)	17.251457	0.002050	0.154974	14.827801
log qtilde(13 , t)	17.227649	0.001034	0.133943	15.724829
log qtilde(14 , t)	16.470473	0.000815	0.105277	15.073375
log qtilde(15 , t)	16.480946	0.000652	0.084473	14.355198
log qtilde(16 , t)	16.379562	0.006606	0.608958	11.461653
log qtilde(17 , t)	16.157130	0.001506	0.095427	13.741173
log qtilde(18 , t)	15.517468	0.001589	0.091209	13.566395
log qtilde(19 , t)	15.798913	0.001026	0.084351	13.889833
log qtilde(20 , t)	15.430008	0.003083	0.101483	12.992467
log qtilde(21 , t)	14.984869	0.010431	0.584183	11.439032
log qtilde(22 , t)	15.077085	0.008350	0.610720	11.115847
log qtilde(23 , t)	15.433582	0.008222	0.586610	11.622434
Clearing Price $\pi(0)$	46.34			

Strike Price	Implied Volatility	Simulated Price	BSM Price	Abs(Sim - BSM)
45.86	2.49%	0.412588	0.531603	0.119015
45.91	2.25%	0.367073	0.494928	0.127854
45.95	2.01%	0.322606	0.459393	0.136787
46.00	1.77%	0.279750	0.425076	0.145326
46.05	1.53%	0.239926	0.392046	0.152120
46.09	1.41%	0.204843	0.360370	0.155527
46.14	0.10%	0.174089	0.330103	0.156015
46.18	0.10%	0.145882	0.301297	0.155415
46.23	4.66%	0.120392	0.273991	0.153599
46.28	6.13%	0.097274	0.248215	0.150941
46.32	6.87%	0.076393	0.223988	0.147595
46.37	7.17%	0.056792	0.201319	0.144527
46.41	7.27%	0.039964	0.180204	0.140240
46.46	7.27%	0.026517	0.160628	0.134111
46.50	7.20%	0.016378	0.142566	0.126188
46.55	7.13%	0.009471	0.125983	0.116512
46.60	7.16%	0.005481	0.110834	0.105353
46.64	7.07%	0.002758	0.097065	0.094307
46.69	6.99%	0.001286	0.084614	0.083328
46.73	6.75%	0.000457	0.073416	0.072959

Table D.2: Simulation summary of European call options on JPM ($\sigma_{\text{BSM}} = 22.16\%$).

Appendix E

Microsoft Corporation (MSFT)

Table E.1: The parameter estimation of Microsoft Corporation (MSFT).

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log Qtilde $p_{\min}$	21.307502	0.000082	0.031389	19.669244
log qtilde(-11 , t)	15.914366	0.000828	0.103554	14.920054
log qtilde(-10 , t)	16.604701	0.001143	0.080203	13.961412
log qtilde(-9 , t)	17.324981	0.001709	0.130279	14.299022
log qtilde(-8 , t)	17.737570	0.000590	0.088344	14.656805
log qtilde(-7 , t)	18.020561	0.000853	0.113889	15.385091
log qtilde(-6 , t)	18.393944	0.000393	0.092662	15.137564
log qtilde(-5 , t)	18.512684	0.000355	0.080933	15.939420
log qtilde(-4 , t)	19.074464	0.000281	0.095195	17.030797
log qtilde(-3 , t)	18.798730	0.000479	0.090834	15.685735
log qtilde(-2 , t)	19.780104	0.000103	0.038112	17.855612
log qtilde(-1 , t)	19.553501	0.000001	0.639663	17.425837
log qtilde(0 , t)	20.030896	0.000098	0.030407	18.610493
log qtilde(1 , t)	20.285612	0.000117	0.047711	18.240503
log qtilde(2 , t)	19.080511	0.000157	0.041404	17.714252
log qtilde(3 , t)	19.721959	0.000139	0.044927	17.868353
log qtilde(4 , t)	18.492731	0.000129	0.038071	17.006981
log qtilde(5 , t)	19.137144	0.000169	0.038352	17.267590
log qtilde(6 , t)	18.316870	0.000241	0.028808	16.872668

Continued on next page

Table E.1 -- *Continued from previous page*

Parameter	**Initial Value**	**Speed Mean Reversion**	**Volatility**	**Long-Term Value**
log qtilde(7 , t)	18.383786	0.000313	0.065769	16.155795
log qtilde(8 , t)	17.841435	0.000516	0.056012	15.781199
log qtilde(9 , t)	17.407306	0.000365	0.036458	15.906892
log qtilde(10 , t)	17.214312	0.005546	0.602412	12.406219
log qtilde(11 , t)	15.942806	0.006031	0.593774	11.821138
log qtilde(12 , t)	14.765218	0.003714	0.561581	11.935753
Clearing Price $\pi(0)$	25.49			

Strike Price	**Implied Volatility**	**Simulated Price**	**BSM Price**	**Abs(Sim - BSM)**
25.23	2.51%	0.226849	0.259415	0.032567
25.26	2.23%	0.201643	0.234244	0.032601
25.28	1.99%	0.176438	0.209128	0.032690
25.31	1.78%	0.151232	0.184138	0.032906
25.33	1.55%	0.126027	0.159409	0.033382
25.36	1.28%	0.100822	0.135170	0.034349
25.38	1.05%	0.075616	0.111771	0.036155
25.41	0.10%	0.050624	0.089681	0.039057
25.43	0.57%	0.030989	0.069450	0.038461
25.46	0.33%	0.020732	0.051623	0.030891
25.48	1.99%	0.014642	0.036631	0.021989
25.51	3.46%	0.010830	0.024684	0.013854
25.53	4.42%	0.008004	0.015719	0.007715
25.56	5.16%	0.005843	0.009419	0.003577
25.58	5.74%	0.004143	0.005291	0.001148
25.61	6.22%	0.002866	0.002776	0.000090
25.63	6.61%	0.001932	0.001357	0.000575
25.66	6.92%	0.001238	0.000616	0.000622
25.68	7.07%	0.000701	0.000260	0.000441
25.71	7.19%	0.000376	0.000101	0.000275

Table E.2: Simulation summary of European call options on MSFT ($\sigma_{\text{BSM}} = 6.17\%$).

Appendix F

Wells Fargo & Co. (WFC)

Table F.1: The parameter estimation of Wells Fargo & Co. (WFC).

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log Qtilde $p_{\min}$	21.610171	0.000106	0.038621	20.214100
log qtilde(-26 , t)	14.683507	0.005742	0.550446	10.294181
log qtilde(-25 , t)	17.165829	0.010828	0.649227	12.034465
log qtilde(-24 , t)	17.541395	0.007681	0.592420	12.297763
log qtilde(-23 , t)	16.343297	0.007272	0.589107	11.457812
log qtilde(-22 , t)	17.419393	0.007834	0.629874	12.212231
log qtilde(-21 , t)	16.081450	0.008031	0.589382	11.274238
log qtilde(-20 , t)	15.307761	0.012552	0.656777	12.639131
log qtilde(-19 , t)	15.507193	0.013022	0.621422	12.214210
log qtilde(-18 , t)	15.665851	0.007656	0.561598	11.380485
log qtilde(-17 , t)	14.365767	0.006417	0.556644	11.582192
log qtilde(-16 , t)	15.779290	0.005764	0.561378	12.107972
log qtilde(-15 , t)	15.822754	0.005200	0.560130	12.150786
log qtilde(-14 , t)	15.566327	0.006119	0.565824	11.620526
log qtilde(-13 , t)	16.404963	0.000964	0.078616	14.484727
log qtilde(-12 , t)	16.738626	0.000497	0.065490	14.649233
log qtilde(-11 , t)	17.401323	0.000306	0.071778	15.879673
log qtilde(-10 , t)	16.900681	0.000365	0.066612	14.696085
log qtilde(-9 , t)	17.491718	0.000395	0.101824	16.796580

Continued on next page

Table F.1 -- *Continued from previous page*

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log qtilde(-8 , t)	18.384323	0.000419	0.066866	16.623075
log qtilde(-7 , t)	18.652994	0.000346	0.071225	17.114443
log qtilde(-6 , t)	18.697276	0.000524	0.080801	17.101981
log qtilde(-5 , t)	19.212301	0.000279	0.063315	17.263715
log qtilde(-4 , t)	19.434829	0.000183	0.052364	17.863194
log qtilde(-3 , t)	19.274731	0.000151	0.050773	17.661279
log qtilde(-2 , t)	19.677786	0.000220	0.063309	18.422304
log qtilde(-1 , t)	20.090085	0.000221	0.056961	18.328439
log qtilde(0 , t)	19.859206	0.000137	0.046920	18.545402
log qtilde(1 , t)	20.060903	0.000119	0.046131	18.486784
log qtilde(2 , t)	20.108231	0.000134	0.043178	17.933101
log qtilde(3 , t)	19.658475	0.000156	0.043878	18.457199
log qtilde(4 , t)	19.351188	0.000410	0.105480	17.890168
log qtilde(5 , t)	19.387413	0.000147	0.035730	18.034323
log qtilde(6 , t)	19.085756	0.001348	0.290429	16.244587
log qtilde(7 , t)	18.564665	0.000183	0.037954	17.652101
log qtilde(8 , t)	18.131843	0.004151	0.632342	14.131854
log qtilde(9 , t)	18.026897	0.000287	0.058444	17.293342
log qtilde(10 , t)	17.206130	0.003818	0.595476	12.242556
log qtilde(11 , t)	16.731057	0.000576	0.091759	16.009787
log qtilde(12 , t)	16.365037	0.001244	0.092121	14.838418
log qtilde(13 , t)	16.460803	0.000510	0.047651	15.663723
log qtilde(14 , t)	16.556040	0.005367	0.596821	12.785502
log qtilde(15 , t)	15.780459	0.001781	0.091748	14.787464
log qtilde(16 , t)	15.596626	0.001575	0.096886	13.840413
log qtilde(17 , t)	16.496012	0.305387	1.034420	14.657109
log qtilde(18 , t)	15.194593	0.007485	0.604412	12.288347
log qtilde(19 , t)	15.908252	0.001476	0.076857	14.327091
log qtilde(20 , t)	15.696256	0.006188	0.578639	10.753409
log qtilde(21 , t)	15.919469	0.010753	0.592152	11.623707
log qtilde(22 , t)	15.370258	0.249898	0.861479	12.375912
log qtilde(23 , t)	14.842162	0.012381	0.673712	11.950697
log qtilde(24 , t)	14.307325	0.007540	0.569014	11.520054

Continued on next page

Table F.1 -- *Continued from previous page*

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log qtilde(25 , t)	11.140966	0.017833	1.357261	8.970546
log qtilde(26 , t)	15.188762	0.147506	0.957799	12.229773
log qtilde(27 , t)	15.681071	0.018426	0.656033	12.626174
Clearing Price $\pi(0)$	32.09			

Strike Price	Implied Volatility	Simulated Price	BSM Price	Abs(Sim - BSM)
31.60	3.52%	0.158711	0.506539	0.347827
31.64	3.29%	0.141175	0.478736	0.337561
31.67	3.06%	0.125481	0.451462	0.325981
31.70	2.83%	0.111348	0.424764	0.313416
31.73	2.60%	0.098092	0.398691	0.300599
31.76	2.37%	0.085687	0.373291	0.287605
31.79	2.14%	0.074119	0.348611	0.274492
31.83	1.91%	0.063488	0.324696	0.261208
31.86	1.91%	0.053747	0.301589	0.247841
31.89	1.45%	0.045219	0.279330	0.234110
31.92	1.23%	0.037484	0.257955	0.220471
31.95	1.00%	0.031056	0.237496	0.206440
31.98	0.77%	0.025846	0.217980	0.192135
32.02	0.10%	0.021364	0.199430	0.178066
32.05	0.10%	0.017435	0.181861	0.164425
32.08	1.04%	0.014384	0.165283	0.150899
32.11	3.00%	0.011736	0.149700	0.137964
32.14	4.16%	0.009439	0.135111	0.125672
32.17	5.06%	0.007447	0.121506	0.114058
32.20	5.81%	0.005803	0.108871	0.103068

Table F.2: Simulation summary of European call options on WFC ($\sigma_{\text{BSM}} = 23.78\%$).

Appendix G

Exxon Mobil Corporation (XOM)

Table G.1: The parameter estimation of Exxon Mobil Corporation (XOM).

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log Qtilde $p_{\min}$	21.009446	0.000874	0.105182	19.366574
log qtilde(-42 , t)	14.624497	0.015907	0.667597	13.253042
log qtilde(-41 , t)	13.086176	0.109015	0.654904	12.784962
log qtilde(-40 , t)	13.517043	1.655877	0.681512	13.456500
log qtilde(-39 , t)	13.546963	0.017421	0.637520	13.422791
log qtilde(-38 , t)	14.337236	0.238144	0.720761	14.150199
log qtilde(-37 , t)	11.406442	0.112032	0.660725	10.909884
log qtilde(-36 , t)	14.778169	0.005571	0.119424	14.307360
log qtilde(-35 , t)	13.771409	0.015736	0.624190	12.643927
log qtilde(-34 , t)	13.900855	0.020482	0.645683	12.541787
log qtilde(-33 , t)	14.840396	0.012766	0.674597	12.452516
log qtilde(-32 , t)	15.858684	0.023343	0.656301	12.748632
log qtilde(-31 , t)	16.316678	0.014694	0.575758	11.319697
log qtilde(-30 , t)	16.621459	0.011320	0.611610	12.090600
log qtilde(-29 , t)	16.085180	0.020683	0.631545	12.319839
log qtilde(-28 , t)	16.403946	0.008498	0.612954	12.342814
log qtilde(-27 , t)	16.794572	0.012565	0.647135	12.965190
log qtilde(-26 , t)	16.015760	0.006862	0.578925	12.250368
log qtilde(-25 , t)	16.564895	0.008096	0.634048	12.061166

Continued on next page

Table G.1 -- *Continued from previous page*

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log qtilde(-24 , t)	16.334819	0.006891	0.587153	12.308659
log qtilde(-23 , t)	15.861266	0.003756	0.526302	10.189219
log qtilde(-22 , t)	16.934737	0.009103	1.067709	14.200639
log qtilde(-21 , t)	16.872026	0.005609	0.587732	12.945562
log qtilde(-20 , t)	16.433677	0.109877	0.994275	14.591879
log qtilde(-19 , t)	16.770052	0.006832	0.636864	13.876517
log qtilde(-18 , t)	16.336445	0.007575	0.615530	12.825321
log qtilde(-17 , t)	17.068033	0.004571	1.075123	15.036913
log qtilde(-16 , t)	17.132050	0.007299	0.670034	14.103101
log qtilde(-15 , t)	16.776926	0.006183	0.616649	12.489031
log qtilde(-14 , t)	17.636969	0.006599	0.685826	12.381801
log qtilde(-13 , t)	17.394973	0.005669	0.632425	14.125453
log qtilde(-12 , t)	16.981143	0.000700	0.121677	15.953401
log qtilde(-11 , t)	17.787079	0.001008	0.127299	15.983494
log qtilde(-10 , t)	17.687728	0.002993	0.552161	13.088659
log qtilde(-9 , t)	17.465590	0.003741	0.592619	15.006547
log qtilde(-8 , t)	17.939001	0.000170	0.038349	16.593822
log qtilde(-7 , t)	18.019855	0.000155	0.040425	16.491533
log qtilde(-6 , t)	17.878104	0.000103	0.027096	16.326497
log qtilde(-5 , t)	18.202410	0.000354	0.098155	16.702483
log qtilde(-4 , t)	17.694646	0.004941	0.636927	11.979650
log qtilde(-3 , t)	18.350818	0.000169	0.056535	16.874324
log qtilde(-2 , t)	18.484851	0.003530	0.590361	13.402573
log qtilde(-1 , t)	17.976992	0.004656	0.668622	14.989766
log qtilde(0 , t)	18.630720	0.000266	0.058779	17.185329
log qtilde(1 , t)	18.652889	0.000208	0.061229	17.046644
log qtilde(2 , t)	17.869626	0.000225	0.048596	16.524917
log qtilde(3 , t)	18.429198	0.000765	0.176275	16.640369
log qtilde(4 , t)	18.348571	0.000539	0.154286	16.736068
log qtilde(5 , t)	17.652008	0.000421	0.090517	16.079576
log qtilde(6 , t)	18.327256	0.000582	0.146176	16.759592
log qtilde(7 , t)	17.732500	0.000000	0.254523	16.230571
log qtilde(8 , t)	18.029448	0.000648	0.133419	16.769615

Continued on next page

Table G.1 -- *Continued from previous page*

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log qtilde(9 , t)	18.049665	0.000598	0.137189	16.881188
log qtilde(10 , t)	17.520974	0.006786	0.642684	13.656752
log qtilde(11 , t)	17.583543	0.000633	0.126360	16.652372
log qtilde(12 , t)	17.831069	0.000598	0.119880	16.480192
log qtilde(13 , t)	16.823282	0.000782	0.107286	15.348908
log qtilde(14 , t)	17.331758	0.000828	0.158687	15.818677
log qtilde(15 , t)	17.420018	0.000533	0.101964	16.417810
log qtilde(16 , t)	16.234774	0.000000	0.192768	15.517698
log qtilde(17 , t)	17.150509	0.000537	0.114283	15.987554
log qtilde(18 , t)	17.052172	0.000700	0.103300	16.210100
log qtilde(19 , t)	15.868872	0.000000	0.136202	15.418991
log qtilde(20 , t)	16.532697	0.000482	0.071984	15.829400
log qtilde(21 , t)	15.898620	0.000657	0.109432	15.229653
log qtilde(22 , t)	16.150773	0.000000	0.147884	15.653435
log qtilde(23 , t)	16.019048	0.005686	0.604533	12.571065
log qtilde(24 , t)	14.990457	0.003187	0.537610	13.366881
log qtilde(25 , t)	15.908980	0.001591	0.127149	13.979765
log qtilde(26 , t)	15.892331	0.005718	0.611074	11.993217
log qtilde(27 , t)	14.966122	0.005526	0.595363	11.957775
log qtilde(28 , t)	15.489474	0.001497	0.101362	14.209992
log qtilde(29 , t)	15.471658	0.000496	0.070639	14.885561
log qtilde(30 , t)	14.349714	0.000000	0.089663	14.637830
log qtilde(31 , t)	14.227637	0.000443	0.067704	14.369477
log qtilde(32 , t)	14.743712	0.013669	0.643578	11.690280
log qtilde(33 , t)	15.197701	0.000881	0.091932	14.130144
log qtilde(34 , t)	14.951405	0.000569	0.095001	14.765100
log qtilde(35 , t)	13.790680	0.006246	0.565805	10.747055
log qtilde(36 , t)	15.323819	0.002893	0.561243	15.128721
log qtilde(37 , t)	14.947941	0.006394	0.607002	11.615039
log qtilde(38 , t)	14.819434	0.006654	0.609638	11.837343
log qtilde(39 , t)	14.820162	0.003236	0.557800	14.643388
log qtilde(40 , t)	14.698835	0.007161	0.606919	11.871582
log qtilde(41 , t)	13.906644	0.009770	0.629712	11.791541

Continued on next page

Table G.1 -- *Continued from previous page*

Parameter	Initial Value	Speed Mean Reversion	Volatility	Long-Term Value
log qtilde(42 , t)	15.040784	0.006759	0.609860	12.217332
log qtilde(43 , t)	14.749103	0.009753	0.633347	12.707569
Clearing Price $\pi(0)$	84.70			

Strike Price	Implied Volatility	Simulated Price	BSM Price	Abs(Sim - BSM)
83.63	3.09%	0.752199	1.370442	0.618243
83.71	3.01%	0.668984	1.310593	0.641609
83.79	2.82%	0.586236	1.252103	0.665867
83.88	2.38%	0.504781	1.195005	0.690224
83.96	2.17%	0.425784	1.139328	0.713544
84.04	1.90%	0.351934	1.085102	0.733168
84.13	1.69%	0.290538	1.032351	0.741813
84.21	1.44%	0.245337	0.981096	0.735760
84.29	1.18%	0.208161	0.931358	0.723197
84.38	0.95%	0.173863	0.883151	0.709288
84.46	0.10%	0.141527	0.836488	0.694961
84.54	0.46%	0.111009	0.791379	0.680369
84.63	2.17%	0.084739	0.747829	0.663090
84.71	3.89%	0.063338	0.705840	0.642503
84.79	4.78%	0.045401	0.665412	0.620012
84.88	5.34%	0.030869	0.626541	0.595671
84.96	5.76%	0.020349	0.589217	0.568868
85.05	6.07%	0.012800	0.553430	0.540630
85.13	6.28%	0.007510	0.519166	0.511656
85.21	6.43%	0.004156	0.486407	0.482252

Table G.2: Simulation summary of European call options on XOM ($\sigma_{\text{BSM}} = 40.23\%$).

Appendix H

Ph.D. Coursework

Tables H.1 and H.2 show 72 units of coursework completed towards the Ph.D. degree.

Course	**Description**
MGT 326	Financial Accounting (*transferred*)
MGT 335	Corporate Finance (*transferred*)
MGT 336	Corporate Governance
MGT 339	Financial Derivatives
MGT 340	Strategy
MGT 373	Financial Strategy and Policy
MGT 402	Asset Management Practicum
MGT 475	Selected Topics in Finance: Fixed Income and Other Investments
MGT 498	Independent Research
MGT 499	Doctoral Study

Table H.1: Ph.D. Management Coursework

Course	Description
MATH 251	Probability
MATH 252	Statistical Inference
MATH 258	Stochastic Processes
MATH 331	Real and Functional Analysis I
MATH 355	Linear Statistical Models
MATH 358	Mathematical Finance
MATH 359	Simulation
MATH 361	Numerical Methods in Finance
MATH 364	Introduction to Scientific Computing
MATH 458B	Optimal Portfolio Theory I
MATH 458C	Optimal Portfolio Theory II
MATH 462	Bayesian Machine Learning
MATH 463	Financial Time Series
MATH 498	Independent Research
MATH 499	Doctoral Study

Table H.2: Ph.D. Mathematical Coursework

Bibliography

Acharya, V. V. and Pedersen, L. (2005). Asset pricing with liquidity risk. *Journal of Financial Economics*, 77:375--410.

Agarwal, A. (2012). High frequency trading: Evolution and the future. *Capgemini's Strategic Analysis Group*, URL http://www.capgemini.com/capitalmarkets.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Economics*, 5:31--56.

Amihud, Y. and Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17:223--249.

Avellaneda, M., Reed, J., and Stoikov, S. (2011). Forecasting prices from level-i quotes in the presence of hidden liquidity. *Algorithmic Finance*, 1(1):35--43.

Bank, P. and Baum, D. (2004). Hedging and portfolio optimization in financial markets with a large trader. *Mathematical Finance*, 14(1):1--18.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9:3--18.

Bhushan, R. (1989). Firm characteristics and analyst following. *Journal of Accounting and Economics*, 11(2):255--274.

Black, F. (1971). Toward a fully automated stock exchange, part i. *Financial Analysts Journal*, 27(4):28--35.

Breen, W. J., Hodrick, L. S., and Korajczyk, R. A. (2002). Predicting equity liquidity. *Management Science*, 48(4):470--483.

Brennan, M. J. and Subrahmanyam, A. (1996). Market microstructure and asset pricing: Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics*, 41:441--464.

Brunnermeier, M. and Pedersen, L. (2009). Market liquidity and funding liquidity. *Review of Financial Studies*, 22(6):2201--2238.

Çetin, U., Jarrow, R., and Protter, P. (2004). Liquidity risk and arbitrage pricing theory. *Finance and Stochastics*, 8(3):311--341.

Çetin, U. and Rogers, L. (2007). Modeling liquidity effects in discrete time. *Mathematical Finance*, 17(1):15--29.

Çetin, U., Soner, H., and Touzi, N. (2010). Option hedging for small investors under liquidity costs. *Finance and Stochastics*, 14(3):317--341.

Chakravarty, S. and Sarkar, A. (1999). Liquidity in u.s. fixed income markets: A comparison of the bid-ask spread in corporate, government and municipal bond markets. *FRB of New York Staff Report*, 73.

Chalmers, J. M. and Kadlec, G. B. (1998). An empirical examination of the amortized spread. *Journal of Financial Economics*, 48:159--188.

Chordia, T., Roll, R., and Subrahmanyam, A. (2001). Market liquidity and trading activity. *The Journal of Finance*, 56(2):501--530.

Cont, R. (2011). Statistical modeling of high-frequency financial data. *Signal Processing Magazine, IEEE*, 28(5):16--25.

Easley, D., Kiefer, N. M., O'Hara, M., and Paperman, J. B. (1996). Liquidity, information, and infrequently traded stocks. *The Journal of Finance*, 51(4):1405--1436.

Eleswarapu, V. R. and Reinganum, M. R. (1993). The seasonal behavior of the liquidity premium in asset pricing. *Journal of Financial Economics*, 34:373--386.

German, D. and Schellhorn, H. (2012). A no-arbitrage model of liquidity in financial markets involving brownian sheets. *SIAM: Society for Industrial and Applied Mathematics*.

Gökay, S. and Soner, H. (2011). Hedging in an illiquid binomial market. *Preprint*.

Gould, M., Porter, M., Williams, S., McDonald, M., Fenn, D., and Howison, S. (2012). Limit order books. *SSRN-id2047160*.

Grossman, S. J. and Miller, M. H. (1988). Liquidity and market structure. *The Journal of Finance*, 43(3):617--633.

Hart, O. (1977). On the profitability of speculation. *Quarterly Journal of Economies*, 91(4):579--597.

Hicks, J. (1962). Liquidity. *The Economic Journal*, 72(288):787--802.

Hilzenrath, D. (2013). High-frequency trading raises concerns at sec. *Washington Post*.

Hirshleifer, J. (1972). Liquidity, uncertainty, and the accumulation of information. *Uncertainty and Expectations in Economics*.

Holmström, B. and Tirole, J. (1993). Market liquidity and performance monitoring. *Journal of Political Economy*, 101(4):678--709.

Holmström, B. and Tirole, J. (1998). Private and public supply of liquidity. *Journal of Political Economy*, 106(1):1--40.

Huang, J. and Wang, J. (2009). Liquidity and market crashes. *The Review of Financial Studies*, 22(7):2607--2643.

Hull, J. C. (2005). *Options, Futures and Other Derivatives*. Prentice Hall, 6th edition.

Jarrow, R. (1992). Market manipulation, bubbles, corners, and short squeezes. *Journal of Financial and Quantitative Analysis*, 27(3):311--336.

Jarrow, R. (1994). Derivative security markets, market manipulation, and option pricing theory. *Journal of Financial and Quantitative Analysis*, 29(2):241--261.

Keynes, J. (1936). *The general theory of employment, interest and money*. ISBN: 978-1-61203-301-3, 2011 Edition. Bottom of the Hill Publishing.

Krylov, N. (2009). On the itô-wentzell formula for distribution-valued processes and related topics. *arXiv:0904.2752v2*.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 53:1315--1335.

O'Hara, M. (2003). Presidential address: Liquidity and price discovery. *The Journal of Finance*, 58(4):1335--1354.

Parlour, C. (1998). Price dynamics in limit order markets. *The Review of Financial Studies*, 11(4):789--816.

Reinganum, M. R. (1981). Misspecification of capital asset pricing - empirical anomalies based on earnings' yields and market values. *Journal of Financial Economics*, 9:19--46.

Rosenberg, B., Reid, K., and Lanstein, R. (1998). Persuasive evidence of market inefficiency. *The Best of The Journal of Portfolio Management*, 48.

Roşu, I. (2009). A dynamic model of the limit order book. *Review of Financial Studies*, 22(11):4601--4641.

Sandås, P. (2001). Adverse selection and competitive market making: Empirical evidence from a limit order market. *The Review of Financial Studies*, 14(3):705--734.

Zheng, B., Moulines, E., and Abergel, F. (2012). Price jump prediction in limit order book. *Quantitative Finance, Trading and Market Microstructure*, available at SSRN 2026454.