University of Tartu

Faculty of Science and Technology
Institute of Mathematics and Statistics

Limit Order Book Modelling

–A Stochastic Approach–

Financial Mathematics
Master's Thesis (15 ECTS)

Author:
Dominic Isaac Andoh

Supervisor:
Assoc. Prof. Raul Kangro

Tartu, 2017

Aktsiaturu tellimusraamatu modelleerimine–Stohhastiline lähenemine–
Magistritöö
Dominic Isaac Andoh

Lühikokkuvõte

Töös rakendatakse ühte tellimusraamatu stohhastilist mudelit AstraZeneca PLC aktsiahinna muutumise kirjeldamiseks päevasisese kauplemise tingimustes. Vaadeldava mudeli korral on võimalik saada mitmetele huvipakkuvatele tellimusraamatu käitumisega setud küsimustele analüütilisel kujul vastused, mis on oluliselt efektiivsem vastuste leidmisest Monte-Carlo simulatsioonide abil. Laplace teisenduse meetodi rakendamise teel saadakse valemid mitmete huvipakkuvate sündmuste tinglike jaotuste jaoks. Töös käsitletakse põhjalikumalt turuhinna kasvamise tõenäosuse arvutamist ning parima ostuhinnaga tehingutellimuse ennem turuhinna muutumist täitmise tõenäosuse leidmist. Võrdlus sündmuste empiiriliste toimumissagedustega näitab, et vaadeldav mudel kirjeldab küllalt hästi vaadeldava aktsia tellimusraamatu käitumise dünaamikat lühikese ajahorisondi korral.
CERCS teaduseriala: P160 Statistika, operatsioonianalüüs, programmeerimine, finants- ja kindlustusmatemaatika.
Märksõnad: Tellimusraamat, Laplace teisendus, päevasisene kauplemine, pidev topeltoksjon.

Limit Order Book Modelling–A Stochastic Approach–
Master's Thesis
Dominic Isaac Andoh

Abstract

We apply a stochastic model to study the continuous-time dynamics of a limit order book for AstraZeneca PLC. The model is analytically tractable and also captures core empirical properties of the order book, which permits us to compute various quantities of interest bypassing the use of simulation. Using the Laplace transform, we are able to compute the conditional distribution of different events given the state of the order book. In this thesis we compute probabilities of increase in the mid-price and probabilities of executing the bid before the mid-price moves. Comparison with empirical frequencies shows that our model captures accurately the short-term dynamics of the limit order book. However, we noted the model is not always applicable due to inconsistencies in the proportionality of cancellation of some order book data.
CERCS research specialisation: P160 Statistics, operations research, programming, actuarial mathematics.
Key Words: Laplace transform, limit order book, continuous double auction.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Equity trading has seen dramatic change in recent years. Traditionally, stocks exchange were organized as floor-based marketplaces where buyers and sellers were represented by intermediaries who arranged the trades between market participants[Tuominen, 2012]. At the moment electronic trading (*order-driven*) has taken over this old system of trading (*quote-driven*). This paradigm shift has been due to the advancement in technology and regulation in securities trading [Kirilenko et al., 2011]. With this electronic trading system, market participants can now develop trading strategies that employ sophiscated algorithms[Tuominen, 2012].

A continuously operating market in which both the buyers and sellers can announce offers to trade at specific quantities and price, and can also initiate trades by accepting such offers is known as *continuous double auction*(CDA). CDA is well-suited to modern financial market and is the basis for almost all electronic trading system. Such electronic trading systems use a *limit order book*, in which unexecuted or partially executed orders are stored whiles awaiting execution [Luckock, 2003].

1.1 Limit Order Book and Market Microstructure

Kakade et al. [2004] gave a vivid description on the behaviour of the limit order book and the market microstructure. Table 1.1 displays a pseudo-data of limit order book, which we use to explain dynamics of the limit order book.

Suppose a trader wishes to purchase 1674 shares of stocks, but he is only willing to buy the 1674 shares at \$144.75 per share or less. He could choose to submit a limit order with

BUY ORDERS		SELL ORDERS	
Price	**Shares**	**Price**	**Shares**
145.50	2245	**145.53**	**433**
145.50	205	**145.55**	**591**
145.25	237	**145.57**	**3925**
145.25	601	145.57	4437
145.25	2000	145.60	4066
144.75	**1674**	145.63	359
144.50	982	145.70	585
144.50	364	145.75	270
144.49	1155	146.00	5272
144.00	759	146.25	2205
143.95	224	146.27	356
143.90	2755	146.50	420

Table 1.1: Pseudo-data of order book

the desired specification, and the order will be placed in a queue called the *buy order book*. In a limit order one specifies both the desired volume (1674 shares) and price. The buy order book is ordered by price, with the highest unexecuted buy price at the top (often referred to as the *bid*). If there are multiple limit orders at the same price, they are ordered by arrival time (with older orders higher in the book). Similarly, a *sell order book* for sell limit orders is maintained this time with the lowest unexecuted sell price on top (referred to as *ask*). For instance a participant would like to sell 433 shares of stocks at \$145.53.

The order book is thus sorted with the most competitive limit orders at the top (high buy prices and low sell prices) down to less competitive limit orders (low buy prices and high sell prices). The bid and ask price are also together referred to as the *inside market*, and the difference between the two is the *spread*. Hence, without the loss of generality the order book is defined to consist exclusively of unexecuted orders in queues waiting to be executed or cancelled.

Anytime a market order arrives, it is immediately matched with the most competitive limit order in the opposite book. Thus, a market order to buy 3500 shares is matched with enough volume on the sell order book to fill the 3500 shares. For instance, in the example of Table 1.1, such an order would be filled by limit sell orders of 433 shares at \$145.53, 591 shares at \$145.55, and then finally 2476 shares at \$145.57. The remaining 1449 shares of this last limit order then remains as the new most competitive of the sell limit order book. Secondly, if a buy (sell) limit order comes in above the ask(below the bid) price, then the order is matched with orders on the opposite book. It is important to note that the prices of executions are the specified prices of limit orders already in the book, not the prices of the incoming order that is immediately executed[Kakade et al., 2004].

1.2 Order Book and Conditional Probability

The order-driven market has been of great interest to market participants because it provides rich data for stochastic modelling. The order book provides intimate information on the process of trading and price formation and reveals the microscopic structure of the market[Bouchaud et al., 2002]. One key motivation for modelling high-frequency dynamics of limit order books is to use the information on the current state to predict its short-term behaviour. The focus is therefore on *conditional probabilities* of events, given the state of the order book[Cont et al., 2010].

Movement of the limit order book in so many ways could be liken to a queuing systems where limit orders wait in a queue to be executed against a market order or cancelled. In this thesis we model a limit order book as a *continuous-time Markov process* which tracks the quantity of limit orders at each price level. This model enables one to calibrate to a high frequency data and reproduce features of the limit order book in a way that is analytically tractable. We then use *Laplace transform techniques* to compute various *conditional probabilities*. These include the probability of the mid-price increasing in the next move, and the probability of executing an order at the bid before the ask qoutes moves. In this thesis we illustrate these computations by modelling from the order book data of AstraZeneca PLC.

In the next chapter we look at related literature and studies that have been conducted on the limit order book. In Chapter 3 we discuss a model for the dynamics of a limit order book, where orders move by independent Poisson precesses. In the chapter that follows (Chapter 4) we describe our data and then estimate the model parameters from our high-frequency order book time series data. Chapter 5 looks at estimating the conditional probailities of interest using the Laplace tranform methods. In the last chapter (Chapter 6) we test the realiability of the proposed model by comparing conditional probabilities computed by Monte Carlo simulation to those computed with the Laplace transforms methods.

Chapter 2

Literature Review

Dynamics and evolution of the limit order book is the core reason for the drive in the continuous auction market. In the past few decades, the theory of finacial market microstructure has attracted myriad of research attention and has seen rapid growth, yet still in an infancy stage due to complexity of the financial market.

One major difficulty of the limit order book is using mathematics to fit an analytically tractable model to the actual data. One usual practice is to use homogeneous Poisson process to simulate the arrival of new orders. Other reasearchers prefer to study the direct effect of traders behaviour on the market structure. This poses alot of challenges since often times there is the need to quantify some qualitative parameters to ensure the effectiveness of the model. Currently what is trending in the financial market is to study the microscopic behaviour of market participant and the aggregate effect it has on the financial market at large. This is very possible due to the availabity of super-rich data provided by the limit order book. In this chapter we take a look at some studies made on limit order book in a continuous auction market. According to differences in modelling methods, the models of limit order book are divided into three types, namely models based on financial economics, models based on econophysics and models based on queueing theory [Jiang et al., 2011].

2.1 Models based on Financial Economics Method

With the assumption that market participants choose at will to submit limit order or market order according to the state of limit order book, Parlour [1998] proposed a more dynamic model for the evolution of limit order book, found optimal order submission strategies depending on the history state of limit order book and expected incoming order flows, and then gave the dynamic characteristics of inflow and outflow of orders in equilibrium.

Hollifield et al. [2004] provided empirical restrictions of a model of optimal order submissions in a limit order market based on the assumption that a trader's optimal order submission depends on the trader's valuation for the asset and the trade-offs between order prices, execution probabilities and picking off risks. Foucault [1999] looked at a model in which strategic traders with different level of tolerance and patience would submit limit orders and observe the dynamics of the limit order book and its ability to rebound to normal behaviour. He afterwards derived the optimal order submission strategy, and found that the proportion of patient and impatient traders and the arrival rates of orders are core factors that determine the dynamics of the limit order book

Basing on the assumption that the driving force of limit order book dynamics come from the competition between liquidity providers and their waiting costs, Rosu [2009] proposed a dynamic model of evolution of the limit order book, without strategic liquidity traders leveraging on information asymmetry, and then studied the price formation mechanism. In this model traders can randomly enter the market, make amendments or cancel outstanding limit orders, and dynamically choose limit order or market order based on the trade-off of transaction price and waiting cost. Then Rosu [2009] showed that the model has an explicit Markov equilibrium where buy (sell) quotes depend only on the number of buy (sell) orders in the limit order book.

Being inspired by Parlour [1998] and Foucault [1999] among others, whose work suggests that limit orders help predict short-term asset return, which contrast earlier studies that implied that informed traders would only make use of market orders [Glosten, 1994, Rock, 1996, Seppi, 1997]; Kozhan and Salmon [2012] examined whether information contained in the limit order book provide economic value in trading scheme. Applying a simple linear model as well as a model-free genetic algorithm that is based on price, order flow and order book information on a Dollar-Sterling tick data, they found that despite the in-sample statistical significance of variables , which describes the structure of the limit order book in tick-by-tick returns, they do not actually add any consistent economic value out-of-sample.

Cohen and Szpruch [2012] combined clasical financial modelling and the limit order book to study the effect of risk-free profit (*latency arbitrage*) on the the limit order book and how to prevent such unbounded profit. They presented a single-moment model for latency arbitrage between two traders in the presence of a limit order book where one trader is assumed to have a significant speed advantage over the other. They found that a fast trader with perfect prior knowledge of the behaviour of the slow trader can manipulate the limit order book to make gain without any risk at all. This profit is however limited as the fast trader cannot force the slow trader to act. They also realized that the fast trader still makes profit even if she has only a prior distribution for the slow trader's actions, but at a certain level of risk.

2.2 Models based on Econophysics Methods

Inspired by similarities between orders in continuous auction market and particles in physical system, physicist in the past fifteen years have been researching on the limit order book in a continuous auction markets. This model where traders are assumed to act randomly is also known as *zero-intelligence* order book modelling, reason being that order arrivals and placement are purely stochastic in nature.

Based on the assumption of independent and identically distributed (IID) stochastic order flow, Smith et al. [2003] developed a microscopic dynamical statistical model for the continuous double auction and analysed it using simulation, dimensional analysis, and theoretical tools based on mean field approximations. Moreover, Smith et al. [2003] found that order size is a more notable determinant factor in limit order market behaviour than tick size in most cases.

Based on empirical regularities in trading order flow in the London Stock Exchange Mike and Farmer [2007] developed a behavioural model for liquidity and volatility in the limit order book. Their studies of order flow uncovered several interesting regularities in the way trading orders are placed and cancelled. Mike and Farmer [2007] showed that this model can reproduce the distributions of return and spread, but out-of-example test is not very satisfactory. Moreover, in their literature Jiang et al. [2011] cited some shortcomings in this model. First, the limit order book to be simulated does not ensure stability and simulation using the empirical parameters may lead to limit order book being emptied by large consecutive market orders. Secondly, the model in Mike and Farmer [2007] did not study volatility clustering which is tackled by Gu and Zhou [2009] in a modified model. Finally, as the authors pointed out, in this model using an exogenous fractional Brownian motion to simulate the long memory of order symbols will result in correlated price returns, which is inconsistent with empirical stylized facts .

To automate the real-time prediction of quantities such as the mid-price movement and price spread crossing, Kercheval and Zhang [2015] proposed a machine learning framework capture the dynamics of high-frequency order book in the equity market. By characterizing each entry in a limit order book with a vector of attributes such as price and volume at different levels, the framework builds a learning model for each quantity of interest with the help of elp of multi-class support vector machines. The proposed framework proved to be effective for short-term price movements forecast.

Liu [2017] proposed a particle-based computational framework based on empirically developed theory of Financial Brownian Particle (FBP) to model the order of trading in the financial market. His methodology enables one to simulate financial prices such as the exchange rate of some given currencies with consideration to different levels of supply and demand imbalances. It also allows one to examine importatnt issues such as the influence bid-ask spread on prices and the distributions of the market orders. His proposed framework yielded results that are in good agreement with published empirical results.

2.3 Models based on Queueing Theory

Cohen et al. [1985], the first to apply queueing theory methods to study the order flow and the limit order book assumed that order arrival rate is constant and buy and sell queues are independent of each other. Also, they obtained the expected order number in the limit order book and waiting time before execution. However, Cohen et al. [1985] has some limitations for not getting the distribution characteristics of relevant indicators [Jiang et al., 2011].

Loosening up assumptions made by Cohen et al. [1985] and applying the theory of queues with preemptive priorities to the problem of two interactive queues, Domowitz and Wang [1994] characterized the structure of the electronic order book in terms of the distributions of the number of buy and sell orders in the system and the waiting time to trade execution. Also, they found that price volatility and trading volume are usually larger, whiles the bid-ask spread is usually smaller in continuous double auction trading mechanism.

Luckock [2003] proposed a model that provides illustration of the way in which microstructural effects can generate price fluctuations, even when the underlying levels of supply and demand remain static and there is no flow of information. Such fluctuations are purely consequences of mismatches between the immediate needs of buyers and sellers, and are described by non- Markov stochastic processes with little resemblance to the more familiar Itô and Lévy processes generally used to model financial data. Luckock [2003] suggested that, in a real market prices are also affected by a variety of economic and business factors unconnected with market microstructure. It is therefore imperative to regard price variations as being made up of two distinct components with quite different stochastic properties: microstructurally generated fluctuations and movements resulting from market forces—that is, from changing levels of supply and demand.

In their work, Cont et al. [2010] proposed a stylized stochastic model describing the dynamics of a limit order book, where the occurrence of different types of events—market orders, limit orders and cancellations—are described in terms of independent Poisson processes. The formulation of the model, which can be viewed as a queuing system, is entirely based on observable quantities and its parameters can be easily estimated from observations of the events in the order book. The model is simple enough to allow analytical computation of various conditional probabilities of order book events via Laplace transform methods, yet rich enough to capture adequately the short-term behavior of the order book. In this thesis we extensively use the approach proposed by Cont et al. [2010] to model the limit order book.

Toke [2015] also studied the analytical properties of a one-sided order book in which the flow of the limit and market orders are Poisson processes and the distribution of cancelled orders is exponential. Using results for the birth-death processes, he built an analytical formula for the shape of a continuous order book which is very close to empirically tested formulas.

Chapter 3

Continuous-time Model for Limit Order Book

The discovery of electronic trading as the major means of trading financial assets has made the order book very pivotal to understanding the dynamics of price formation. Buy and sell orders are matched continuously in an order-driven market and subject to price and time prioity. *Order book* is a list of buy and sell limit orders, with their corresponding prices and size at a particular time instant.

3.1 The Order Book

Market participant basically can submit three types of orders, namely: *Limt order, Market order and Cancellation order.* Particpants exercise limit order when they quote the price at which they are willing to buy or sell a fix amount of shares. The least price for which there is an outstanding limit sell order is called *ask price* whereas the highest buy price is called *bid price*

Participants exercise a market order when they buy or sell a certain number of shares at the best available opposite quote. Once a market order has arrived it is matched with the best available price in the order book and trade occurs. The quantities available in the limit order book are then updated accordingly. Limit orders remain in the order book until they are executed against a market order or cancelled. Participants who submit exclusively limit orders are referred to as *liquidity providers* whereas those who submit market orders are referred to as *liquidity takers*.

Similar to the work of Cont et al. [2010], in this thesis we assume a market where limit orders are placed on a price grid $\{1, \cdots, n\}$ representing multiples of a price tick. The state of the order book is monitored with a continuous-time process $X(t) \equiv (X_1(t), \cdots, X_n(t))_{t \geq 0}$, where $|X_p(t)|$ is the number of outstanding limit orders at price p, where $1 \leq p \leq n$. If $X_p(t) < 0$ then there are $-X_p(t)$ bid orders at price p; if $X_p(t) > 0$, then there are $X_p(t)$ ask orders at price p. We let the *ask price* $p_A(t)$ at time t be:

$$p_A(t) = \min(\inf\{p = 1, \cdots, n, X_p(t) > 0\}, (n+1))$$

and the *bid price* $p_B(t)$ at time t as:

$$p_B(t) = \max(\sup\{p = 1, \cdots, n, X_p(t) < 0\}, 0)$$

In situations where there are no ask orders in the order book an ask price of $n+1$ is forced and when there are no bid orders a bid price of 0 is forced. The *mid-price* $p_M(t)$ and the *bid-ask spread* $s(t)$ are defined as:

$$p_M(t) \equiv \tfrac{p_B(t)+p_A(t)}{2} \qquad \text{and} \qquad s(t) \equiv p_A(t) - p_B(t)$$

According to Bouchaud et al. [2002], most market activities occur in the neighbourhood of the bid and ask prices and it is therefore essential to keep track of the number of orders at a given distance from the bid/ask. In this light we define the number of buy orders at a distance i from the ask as:

$$Q_i^B(t) = \begin{cases} X_{p_A(t)-i}(t) & \text{where} \quad 0 < i < p_A(t) \\ 0 & \text{where} \quad p_A(t) \leq i < n \end{cases} \tag{3.1}$$

and the number of sell orders at a distance i from the bid as:

$$Q_i^A(t) = \begin{cases} X_{p_B(t)+i}(t) & \text{where} \quad 0 < i < p_B(t) \\ 0 & \text{where} \quad n - p_B(t) \leq i < n \end{cases} \tag{3.2}$$

$Q_i^A(t)$ and $Q_i^B(t)$ throw more light on the on the shape of the order book relative to the best opposite quotes [Cont et al., 2010].

3.2 Dynamics of the Order Book

In this section we describe the model as Cont et al. [2010] proposed in their work, This has got to do with how the order book is updated by the influx of new orders. Given a state $x \in \mathbb{Z}^n$ and $1 \leq p \leq n$, we define :

$$e_p \equiv (0, \cdots, 1, \cdots, 0),$$

where 1 in the vector e_p is in the p^{th} component. It is assumed that all orders are of unit size. For empirical example in this thesis, we take the unit size to be the average size of limit orders observed for the asset. In their 2010 work Cont et al. [2010] outlined how the inflow of different orders affect the shape of the order book:

- a limit buy order at the price level $p < p_A(t)$ increases the quantity at level $p : x \to x - e_p$
- a limit sell order at the price level $p > p_B(t)$ increases the quantity at level $p : x \to x + e_p$
- a market buy order decreases the quantity at the ask price $x \to x - e_{p_A(t)}$
- a market sell order decreases the quantity at the ask price $x \to x + e_{p_B(t)}$
- a cancellation of an outstanding limit buy order at price level $p < p_A(t)$ decreases the quantity at level $p : x \to x + e_p$
- a cancellation of an oustanding limit sell order at price level $p > p_B(t)$ decreases the quantity at level $p : x \to x - e_p$

The evolution of the order book is therefore an interplay between the order book and the order flow.

Bouchaud et al. [2002] observed that incoming orders (limit, market and cancellation orders) arrive more frequently in the vicinity of the current bid/ask price and the arrival rate of these orders depend on the distance to the bid/ask. In order to capture these empirical features in an analytically tractable manner that can allow computation of quantities of interest –especially the conditional probabilities of different events– we use the stochastic model proposed by Cont et al. [2010] where the above outlined events are simulated using independent Poisson process, assuming that $\forall i \geq 1$:

- Limit buy (sell) orders arrive at a distance of i ticks from the opposite best quote at independent, exponential times with the rate $\lambda(i)$,

- Market buy (sell) orders arrive at independent, exponential times with rate μ,

- Cancellations of limit orders at a distance of i ticks from the opposite best quote occur at a rate proportional to the number of outstanding orders: hence given that the number of outstanding orders at that level is x then the cancellation rate is $\theta(i)x$.

- The above events are mutually independent.

With these above assumptions one could say that X is a continuous-time Markov chain with a state space $\mathbb{Z}^n$ and transition rates given by the following:

$$
\begin{gathered}
x \to x - e_p \text{ with rate } \lambda(p_A(t) - p) \text{ for } p < p_A(t),\\
x \to x + e_p \text{ with rate } \lambda(p - p_B(t)) \text{ for } p > p_B(t),\\
x \to x - e_{p_B(t)} \text{ with rate } \mu,\\
x \to x - e_{p_A(t)} \text{ with rate } \mu,\\
x \to x + e_p \text{ with rate } \theta(p_A(t) - p)|x_p| \text{ for } p < p_A(t),\\
x \to x - e_p \text{ with rate } \theta(p - p_B(t))|x_p| \text{ for } p > p_B(t),
\end{gathered}
$$

Chapter 4

Parameter Estimation

4.1 Data Set Description

In this thesis we use a 2006 order book for AstraZeneca PLC which consist of timed-stamped sequence of trades (market order) and quotes (prices, quantities of outstanding limit orders) for the 10 best price levels on each side of the order book.

In Table 4.1, we show five consecutive trades of AstraZeneca PLC. Each row shows bid1 and ask1 where trade occurs together with their respective sizes on each price level and also the time of trade. Table 4.2 displays ask side qoutes. Each row displays 5 out of the 10 ask prices (ask1,ask2,. . . ,ask5), as well as the quantity of shares asked at their respective prices (asize1,asize2,. . . ,asize5). This is true for the bid side of the book as well.

	time	price	bid1	bsize1	ask1	asize1
1	08:30:03.316	2853	2852	2820	2853	8288
2	08:30:03.316	2853	2852	2820	2853	2496
3	08:30:10.075	2852	2852	944	2853	8704
4	08:30:52.093	2852	2851	5214	2853	13874
5	08:30:59.603	2851	2851	4126	2853	14474

Table 4.1: A sample of five trades for AstraZeneca PLC

	ask1	ask2	ask3	ask4	ask5	asize1	asize2	asize3	asize4	asize5
1	2853	2854	2855	2857	2858	8288	1600	22488	6283	5876
2	2853	2854	2855	2857	2858	2496	1600	22488	6283	5876
3	2853	2854	2855	2857	2858	1896	1600	22488	6283	5876
4	2853	2854	2855	2857	2858	8704	1600	22488	6283	5876
5	2853	2854	2855	2857	2858	8704	1600	22488	6283	5876

Table 4.2: A sample of five qoutes for AstraZeneca PLC

4.2 Estimation Procedure

From the data set, we first compute the average size of market orders S_m, the average size of the limit orders S_l and also the average size of the cancelled orders S_c. We then make the average size of the limit order S_l the unit size. Thus a block of orders of size S_l is counted as one event. The limit order arrival rate function for $1 \leq i \leq 10$ is estimated by:

$$\hat{\lambda}(i) = \frac{N_l(i)}{T},$$

where $N_l(i)$ is the total number of limit orders that arrived at a distance i from the opposite best quote. $N_l(i)$ is obtained by counting the number of times a quote increases in size at a distance of $1 \leq i \leq 10$ ticks from the opposite best quote, $\hat{\lambda}(i)$ is then extrapolated by fitting a power law function

$$\hat{\lambda}(i) = \frac{k}{i^\alpha}$$

as propsed by Bouchaud et al. [2002]. The parameters of the power law parameters k and α are obtained by a least square fit:

$$\min_{k,\alpha} \sum_{i=1}^{10} \left(\hat{\lambda}(i) - \frac{k}{i^\alpha} \right)^2$$

Figure 4.1 on page 15 displays the estimated arrival rate at distance $0 \leq i \leq 20$ from the opposite best quotes. We realized that this approach gave a bad fit. In order to improve the fitted model to our data we still used the approach by Bouchaud et al. [2002], but this time using the first two limit order rates as they appear and then fit a curve tstarting from a distance $3 \leq i \leq 20$ from the opposite best quotes. Figure 4.2 shows these improved fitted limit order rates.

Arrival rates of the market orders is estimated by:

$$\hat{\mu} = \frac{N_m}{T} \frac{S_m}{S_l},$$

where T is the length of time (one full working day) in minutes and N_m is the number of market orders.

The cancellation rate in this model is proportional to the number of orders at a particular

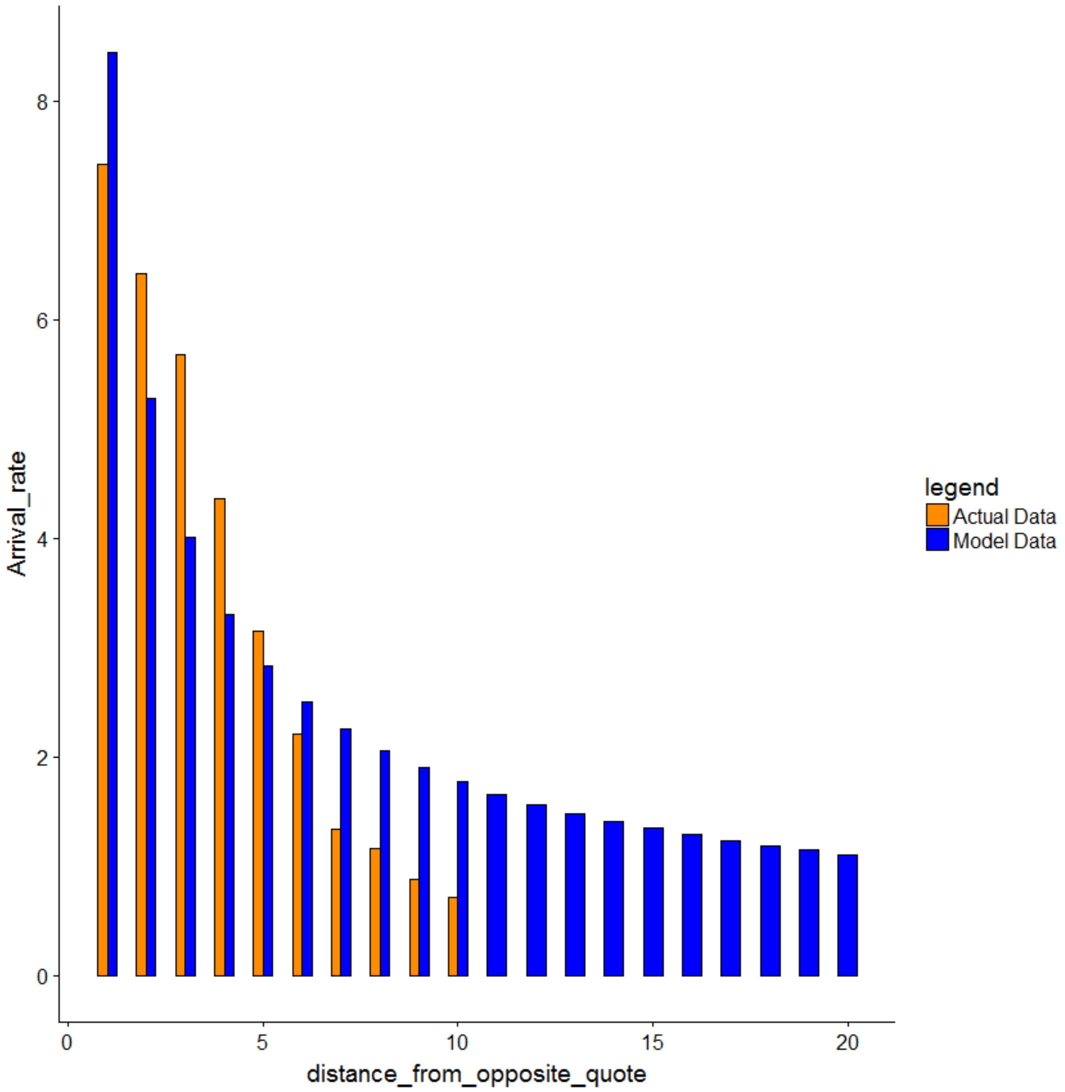

Figure 4.1: Fitted limit order rates

price level. Therefore one approach is to estimate the steady state shape of the order book Q_i when estimating the cancellation rate, which is given as the average number of orders at a distance of i ticks from the opposite best quote, for $1 \leq i \leq 10$. Given that M is the number of quote rows and $S_i^B(j)$ the number of shares bid at a distance of i ticks from the ask on the j^{th} row for all $1 \leq j \leq M$, we have:

$$Q_i^B = \frac{1}{S_l}\frac{1}{M}\sum_{j=1}^{M} S_i^B(j)$$

The vector Q_i^A is obtained in a similar fashion and Q_i is the average of Q_i^A and Q_i^B. The cancellation rate function is estimated by:

$$\hat{\theta}(i) = \frac{N_c(i)}{TQ_i}\frac{S_c}{S_l} \quad \text{for} \quad i \leq 10 \quad \text{and} \quad \hat{\theta}(i) = \hat{\theta}(10) \quad \text{for} \quad i > 10$$

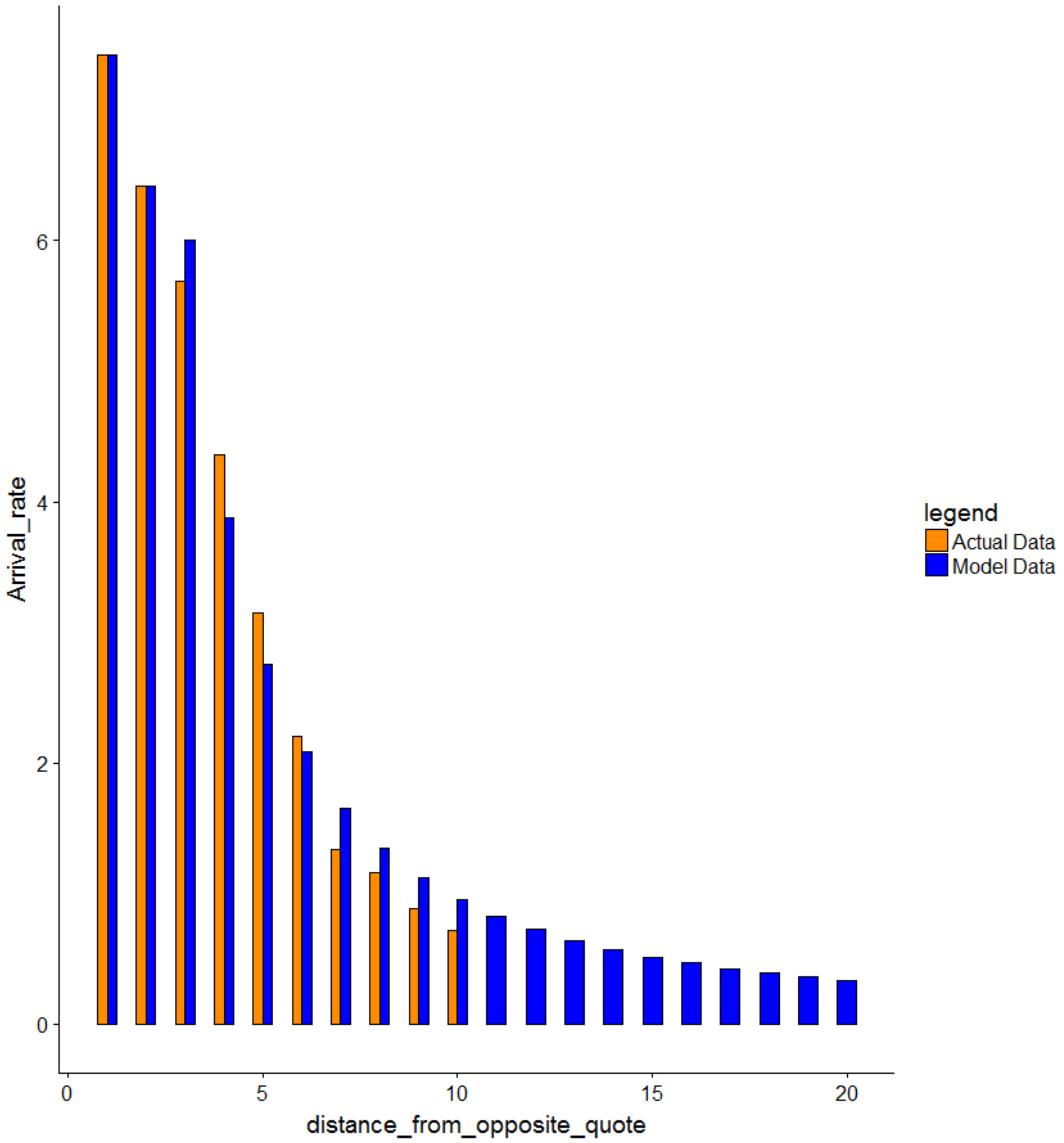

Figure 4.2: Improved fitted limit order rates

We extend the values of the cancellation rates by a constant and they are displayed by Figure 4.3 on page 17. Again $N_c(i)$ is obtained by counting the number of times a quote diminishes in size at a distance $1 \leq i \leq 10$ ticks from the opposite best quote, not counting decreases that are due to market orders.

The estimated parameter values for AstraZeneca PLC are given in Table 4.3.

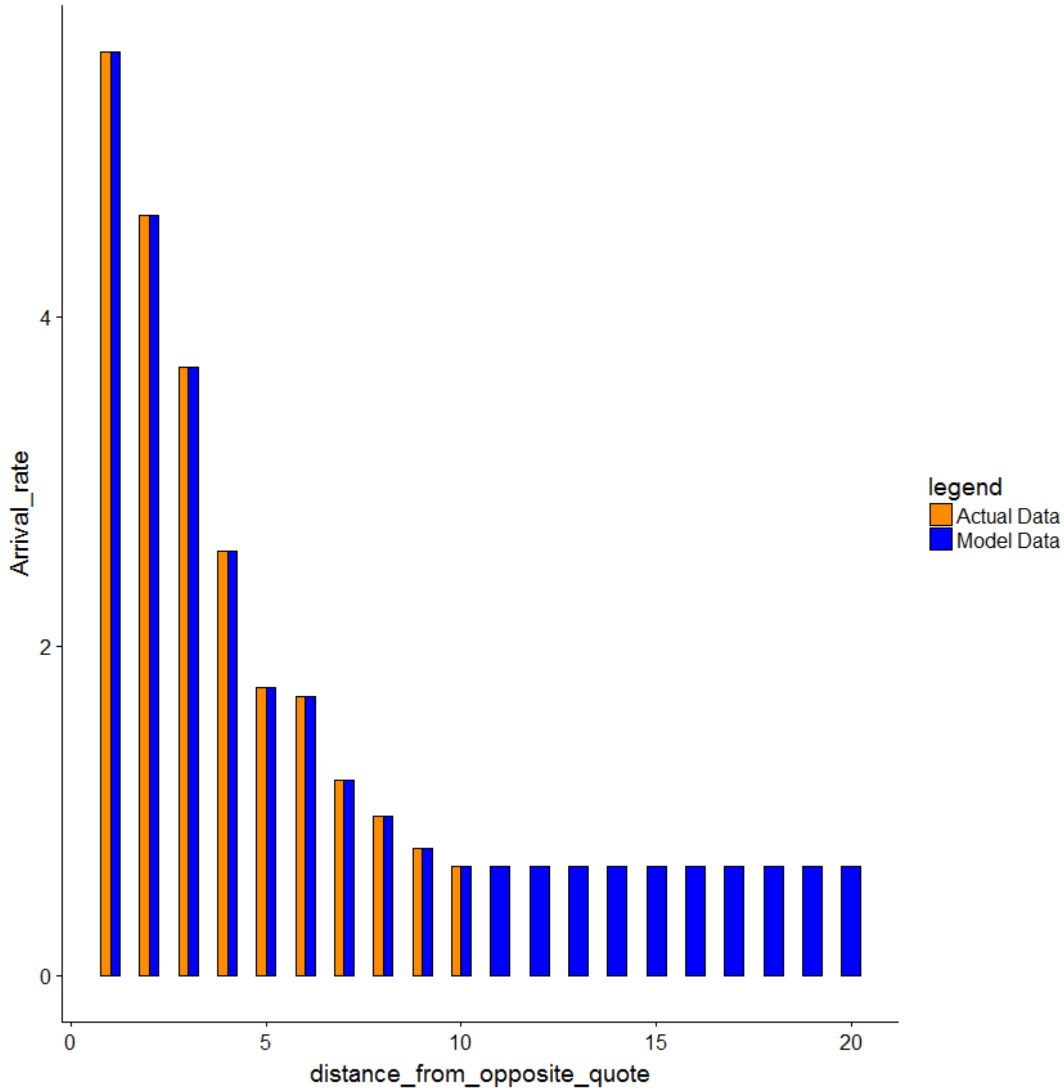

Figure 4.3: Extended cancellation rates

i	1	2	3	4	5	6	7	8	9	10
$\hat{\lambda}(i)$	7.43	6.42	5.69	4.36	3.15	2.20	1.33	1.16	0.88	0.72
$\hat{\theta}(i)$	5.61	4.61	3.69	2.58	1.75	1.69	1.18	0.97	0.77	0.66
$\hat{\mu}$	7.18									
k	8.45									
α	0.68									

Table 4.3: Estimated parameters for AstraZeneca PLC

Chapter 5

Laplace Transform & Conditional Probability

A key motivation for modelling a high-frequency dynamics of the order book is to use the deduced information to predict the short-term behaviour of various quantities of interest that are useful to traders and also useful in algorithmic trading. In this thesis we look at the probability of the mid-price moving up and the probability of executing a limit order at the bid before the ask quote moves. Here we show that the Laplace transform method could be used to compute the conditional probabilities of these quanties bypassing the need for Monte Carlo simulation.

5.1 Birth-Death Processes

Definition 5.1.1. *Let $f(t)$ be defined for $t \geq 0$. The Laplace transform of $f(t)$, denoted by $\hat{f}(s)$, is an integral transform given by the Laplace integral:*

$$\hat{f}(s) = \int_0^\infty e^{-st} f(t) dt$$

where s is a complex number.

The two-sided Laplace transform allows time function to be non-zero for negative time and is given by :

$$\hat{f}(s) = \int_{-\infty}^\infty e^{-st} f(t) dt \tag{5.1}$$

Suppose f is a probability density function (pdf) of some random variable X, then the two-sided Laplace transform can also be represented as:

$$\hat{f}(s) = \mathbb{E}[e^{-sX}] \tag{5.2}$$

In this case, we also say that $\hat{f}$ is the two-sided Laplace transform of the random variable itself [Cont et al., 2010]. We use the *two-sided* Laplace transform in this thesis because our function f follows a pdf of random walk with both negative and positive (ask and bid) sides. Going forward, we drop the prefix "two-sided" when referring to two-sided Laplace transforms. For instance we imply the Laplace transform of the conditional pdf of X given A, when we say conditional Laplace transform of the random variable X conditional on the event A.

Given that X and Y are independent random variables with a well-defined Laplace transforms, then by Equation 5.2 we have:

$$\hat{f}_{X+Y}(s) = \mathbb{E}[e^{-s(X+Y)}] = \mathbb{E}[e^{-sX}]\mathbb{E}[e^{-sY}] = \hat{f}_X(s)\hat{f}_Y(s). \tag{5.3}$$

If for some $\gamma \in \mathbb{R}$ we have $\int_{-\infty}^{\infty} |\hat{f}(\gamma + i\omega)| d\omega < \infty$ and $f(t)$ is a continuous function of t, then the *inverse* Laplace transform formula given by Bromwich integral is

$$f(t) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{ts} \hat{f}(s) ds. \tag{5.4}$$

Consider a birth-death process with constant birth rate λ and death rate μ_i $\forall i \geq 1$ and let σ_b be the first passage time of the process from the state b to 0 . Let's assume σ_b can be represented as:

$$\sigma_b = \sigma_{b,b-1} + \sigma_{b-1,b-2} + \cdots + \sigma_{1,0} \tag{5.5}$$

where all terms on the RHS are independent. In addition, suppose $\sigma_{i,i-1}$ denotes the first-passage time of the birth-death process from state i to $i-1$ $\forall i \in [1, b]$. Let $\hat{f}_b$ denotes the Laplace transform of σ_b and $\hat{f}_{i,i-1}$ the Laplace transform of $\sigma_{i,i-1} \forall i \in [1, b]$ then from Equation 5.3 we can write:

$$\hat{f}_b(s) = \prod_{i=1}^{b} \hat{f}_{i,i-1}(s) \tag{5.6}$$

Since computing the Laplace transform of $\hat{f}_{i,i-1} \forall i \in [1, b]$ is simpler than computing that of $\hat{f}_b$, it suffice to compute the simpler $\hat{f}_{i,i-1} \forall i \in [1, b]$.

5.2 A Conditional Exponential Probability

Lemma 5.2.1. *Assume $T_1 \sim Exp(\lambda)$ and $T_2 \sim Exp(\mu)$ are independent random variables. Then*

$$\mathbf{Pr}(T_2 < T_1) = \frac{\mu}{\lambda + \mu}$$

and the conditional distribution of T_2, under the condition $T_2 < T_1$ is given by

$$T_2 | T_2 < T_1 \sim Exp(\lambda + \mu).$$

Proof. Let us find the cdf of $T_2 | T_2 < T_1$. According to the definition of cdf and the definition of conditional probability, we have

$$\begin{aligned} F_{T_2|T_2<T_1}(t) &\overset{\mathbf{def}}{=} \mathbf{Pr}(T_2 \le t | T_2 < T_1) \\ &= \frac{\mathbf{Pr}(T_2 \le t, T_2 < T_1)}{\mathbf{Pr}(T_2 < T_1)} \end{aligned}$$

Now

$$\mathbf{Pr}(T_2 \le t, T_2 < T_1) = \mathbf{Pr}((T_1, T_2) \in \mathcal{D}),$$

where $\mathcal{D} = \{(t_1, t_2) \quad \text{such that} \quad 0 \le t_2 \le t \quad \text{and} \quad t_1 \ge t_2\}$ as displayed in the shaded area in Figure 5.1. The probability is given by the double integral over the area $\mathcal{D}$:

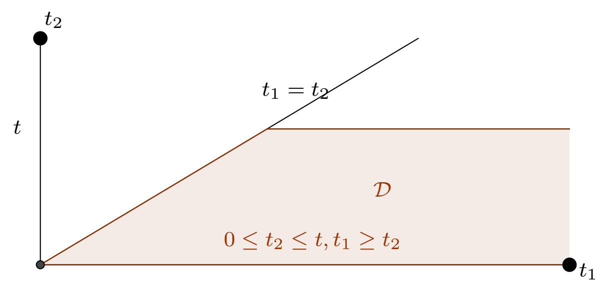

Figure 5.1: Area of Probability

$$
\begin{aligned}
\mathbf{Pr}[(T_1, T_2) \in \mathcal{D}] &= \iint_{\mathcal{D}} f_{T_1}(t_1) f_{T_2}(t_2)\, dt_1\, dt_2 \\
&= \int_0^t \left(\int_{t_2}^{\infty} \lambda e^{-\lambda t_1} \mu e^{-\mu t_2}\, dt_1 \right) dt_2 \\
&= \int_0^t \mu e^{-\mu t_2} \left(\int_{t_2}^{\infty} \lambda e^{-\lambda t_1}\, dt_1 \right) dt_2 \\
&= \int_0^t \mu e^{-\mu t_2} e^{-\lambda t_2}\, dt_2 \\
&= \int_0^t \mu e^{-(\lambda+\mu) t_2}\, dt_2 \\
&= \frac{\mu}{\lambda+\mu}(1 - e^{-(\lambda+\mu)t}).
\end{aligned}
$$

since

$$
\mathbf{Pr}(T_2 < T_1) = \lim_{t \to \infty} \mathbf{Pr}(T_2 \leq t, T_2 < T_1), \tag{5.7}
$$

we have shown that

$$
\mathbf{Pr}(T_2 \leq T_1) = \frac{\mu}{\lambda+\mu} \tag{5.8}
$$

and

$$
F_{T_2|T_2<T_1}(t) = \begin{cases} 1 - e^{-(\lambda+\mu)} & \text{where} \quad t \geq 0 \\ 0 & \text{where} \quad t < 0 \end{cases}. \tag{5.9}
$$

Note that $F_{T_2|T_2<T_1}(t)$ corresponds to the exponential distribution $Exp(\lambda+\mu)$ therefore

$$
\Rightarrow T_2 | T_2 < T_1 \sim Exp(\lambda+\mu) \quad \text{and} \quad \mathbb{E}[e^{-sT}] = \frac{\lambda}{\lambda+s} \quad \text{if} \quad T \sim Exp(\lambda).
$$

□

Given that a birth-death process has a constant birth rate λ and a death rate μ_i, then from the Formula (5.8) in the proof of Lemma 5.2.1 we can say that in the first transition from the state i, the process moves down to state $i-1$ with probability $\mu_i/(\lambda+\mu_i)$. It moves up to state $i+1$ with the prbability $\lambda/(\lambda+\mu_i)$. The first transition state could happen in either of these two ways; first the process could move down form i to $i-1$, if the process moves down the first passage is complete, or it could move up to $i+1$ then it must move from $i+1$ to i and then move down from i to $i-1$. Let $T_1 \sim Exp(\lambda)$ be the time to the next birth and $T_2 \sim Exp(\mu_i)$ be the time to the next death, then according to the law of *Total Expectation* we have

$$\begin{aligned}\hat{f}_{i,i-1}(s) &= \mathbb{E}[e^{-sT_{i,i-1}}] \quad \text{from Lemma 5.2.1 we know that } T_1, T_2 \text{ are exponential, hence we write} \\ &= \mathbf{Pr}(T_2 < T_1)\mathbb{E}[e^{-sT_2)}|T_2 < T_1] + \mathbf{Pr}(T_1 \leq T_2)\mathbb{E}[e^{-s(T_1+\sigma_{i+1,i}+\sigma_{i,i-1})}|T_1 \leq T_2].\end{aligned}$$

Let T be a random variable from $Exp(\lambda + \mu_i)$, then according to the Lemma 5.2.1 we have

$$\mathbb{E}[e^{-sT_2)}|T_2 < T_1] = \mathbb{E}[e^{-sT}]$$

and

$$\mathbb{E}[e^{-sT_1)}|T_1 \leq T_2] = \mathbb{E}[e^{-sT}].$$

Thus

$$\hat{f}_{i,i-1}(s) = \frac{\mu_i}{\lambda + \mu_i}\mathbb{E}[e^{sT}] + \frac{\lambda}{\lambda + \mu_i}\left[\mathbb{E}[e^{sT}]\hat{f}_{i+1,i}(s)\hat{f}_{i,i-1}(s)\right]. \tag{5.10}$$

Since

$$\mathbb{E}[e^{-sT}] = \int_0^\infty (\lambda + \mu_i)e^{-(\lambda+\mu_i+s)x}dx = \frac{\lambda + \mu_i}{\lambda + \mu_i + s},$$

we finally get

$$\begin{aligned}\hat{f}_{i,i-1}(s) &= \left(\frac{\mu_i}{\lambda + \mu_i}\right)\left(\frac{\lambda + \mu_i}{\lambda + \mu_i + s}\right) + \left(\frac{\lambda}{\lambda + \mu_i}\right)\left[\left(\frac{\lambda + \mu_i}{\lambda + \mu_i + s}\right)\hat{f}_{i+1,i}(s)\hat{f}_{i,i-1}(s)\right] \\ &= \frac{\mu_i}{\lambda + \mu_i + s} + \frac{\lambda\hat{f}_{i+1,i}(s)\hat{f}_{i,i-1}(s)}{\lambda + \mu_i + s}.\end{aligned}$$

Solving this equation for $\hat{f}_{i,i-1}(s)$, we get

$$\hat{f}_{i,i-1}(s) = \frac{\mu_i}{\lambda + \mu_i + s - \lambda\hat{f}_{i+1,i}(s)} \tag{5.11}$$

Iterating on quantity 5.11 produces

$$\hat{f}_{i,i-1}(s) = -\frac{1}{\lambda}\mathbf{\Phi}_{\mathbf{k=i}}^{\infty}\frac{-\lambda\mu_k}{\lambda + \mu_k + s} \tag{5.12}$$

where the term on the RHS is the corresponding generalized *continuous fraction*. Combining Equations 5.6 and 5.12 give us

$$\hat{f}_b(s) = \left(-\frac{1}{\lambda}\right)^b \left(\prod_{i=1}^{b} \mathbf{\Phi}_{\mathbf{k=i}}^{\infty}\frac{-\lambda\mu_k}{\lambda + \mu_k + s}\right) \tag{5.13}$$

Quantity 5.13 would be use to compute probabilities of interest in this thesis.

Proposition 5.2.2. : ***Laplace transform of Cumulative Distribution Function (cdf)***
Suppose X is an exponential random variable with the pdf $f_X(x)$ which satisfies the condition $\lim_{s\to-\infty} f_X(x)/e^{s_0x} = 0$ for some $s_0 > 0$, and a cdf

$$F_X(x) = \int_{-\infty}^{x} f_X(z)\,dz,$$

then the Laplace transform given by

$$\begin{aligned}\mathfrak{L}[F_X(s)] &= \frac{1}{s}\int_{-\infty}^{\infty} e^{-sx} f_X(x)dx &(5.14)\\ &= \frac{1}{s}\hat{f}_X(s) &(5.15)\end{aligned}$$

given that $0 < s \leq s_0$

Proof. Let

$$\mathfrak{L}[F_X(s)] = \int_{-\infty}^{\infty} e^{-sx} F_X(x)dx \tag{5.16}$$

we use integration by parts to simplify Equation 5.16. We make $u = F_X(x)$ such that $u' = f_X(x)$, again we let $v' = e^{-sx}$ such that $v = -\frac{1}{s}e^{-sx}$ then applying integration parts gives us

$$\int_{-\infty}^{\infty} e^{-sx} F_X(x)dx = -\frac{1}{s}\frac{F_X(x)}{e^{sx}}\bigg|_{-\infty}^{\infty} + \frac{1}{s}\int_{-\infty}^{\infty} e^{-sx} f_X(x)dx. \tag{5.17}$$

Applying the *L'Hospital's rule* of indeterminate forms on LHS of Equation 5.17 causes the first term to run to zero for $0 < s \leq s_0$, leaving the equation as

$$\begin{aligned}\int_{-\infty}^{\infty} e^{-sx} F_X(x)dx &= \frac{1}{s}\int_{-\infty}^{\infty} e^{-sx} f_X(x)dx\\ &= \frac{1}{s}\hat{f}_X(s)\end{aligned}$$

□

5.3 Direction of Mid-price Move

At this section we look at computing the probability for an increase in the mid-price at the next move. The mid-price first moves when the bid or ask queue goes to zero at the first-passage time or, if the bid/ask spread is greater than one the first time a limit order enters the spread. In this section we let X_A denote the average size of the best limit sell order (ask) and X_B the average size of the best buy order (bid). We let σ_A and σ_B be the first-passage times of X_A and X_B to 0 respectively. Again we let $W_B(t)$ (W_A(t)) $\forall t \geq 0$ denote the number of outstanding orders at the bid (ask) at time t of the $X_B(0)$ ($X_A(0)$). We make ϵ_B (ϵ_A) the first-passage time for W_B (W_A) to 0. Finally, we let T be the time of the first change in the mid-price such that:

$$T \equiv \inf\{t \geq 0, p_M(t) \neq p_M(0)\}.$$

Our interest here is to compute the conditional probability that the mid-price will increase before decreasing:

$$\mathbf{Pr}[p_M(T) > p_M(0) | X_A(0) = a, X_B(0) = b, s(0) = S], \tag{5.18}$$

where $S > 0$.

The motive of computing (5.18) is to observe that X_A and X_B behave as independent birth-death processes whereas W_A and W_B behave as independent pure-death processes for $t \leq T$ [Cont et al., 2010]. To expantiate on this claim Cont et al. [2010] outlined the following lemma in their work:

Lemma 5.3.1. *Let s(0)=S, then*

1. *There exist independent birth-death processes $\tilde{X}_A$ and $\tilde{X}_B$ with each having birth rate $\lambda(S)/2$ and death rate $(\mu + i\theta(S))/2$ in state $i \geq 1$, such that $\forall t \in [0, T]$, $\tilde{X}_A(t) = X_A(t)$ and $\tilde{X}_B(t) = X_B(t)$.*

2. *There exist independent pure death processes $\tilde{W}_A$ and $\tilde{W}_B$ with each having death rate $(\mu + i\theta(S))/2$ in state $i \geq 1$, such that $\forall t \in [0, T]$, $\tilde{W}_A = W_A(t)$ and $\tilde{W}_B(t) = W_B(t)$. Furthermore, $\tilde{W}_A$ is independent of $\tilde{X}_B$, $\tilde{W}_B$ is independent of $\tilde{X}_A$, $\tilde{W}_A \leq \tilde{X}_A$ and $\tilde{W}_B \leq \tilde{X}_B$.*

We compute quantity 5.18 as follows:

Proposition 5.3.2. : *Probability of increase in the mid-price*

Given that

$$\hat{f}_{j,S}(s) = \left(-\frac{1}{\lambda(S)/2}\right)^j \left(\prod_{i=1}^{j} \mathbf{\Phi}_{\mathbf{k=i}}^{\infty} \frac{-\lambda(S)/2(\mu + k\theta(S))/2}{(\lambda(S) + \mu + k\theta(S))/2 + s}\right), \tag{5.19}$$

if $j \geq 1$ and $\Lambda_S \equiv \sum_{i=1}^{S-1} \lambda(i)/2$, then (5.18) is given by the inverse Laplace transform (Equation 5.4 at page 19) of:

$$\hat{F}_{a,b,S}(s) = \frac{1}{s}\left(\hat{f}_{a,S}(s+\Lambda_S) + \frac{\Lambda_S}{\Lambda_S + s}(1 - \hat{f}_{a,S}(s+\Lambda_S))\right)\left(\hat{f}_{b,S}(\Lambda_S - s) + \frac{\Lambda_S}{\Lambda_S - s}(1 - \hat{f}_{b,S}(\Lambda_S - s))\right), \tag{5.20}$$

evaluated at 0. *When $S = 1$ (5.20) reduces to*

$$\hat{F}_{a,b,1}(s) = \frac{1}{s}\hat{f}_{a,1}(s)\hat{f}_{b,1}(-s). \tag{5.21}$$

Proof. In this thesis we are interested in computing the conditional probability of mid-price increase in the special case when $S = 1$ and therefore the *proof* would be base on this special case only.

We construct $\tilde{X}_A$ and $\tilde{X}_B$ as indicated in Lemma 5.3.1. In the case when $S = 1$, the price changes for the first time exactly at the point when one of the two independent birth-death processes $\tilde{X}_A$ and $\tilde{X}_B$ empties to 0 for the first time. Each of these birth-death processes has a constant birth rate of $\lambda(1)/2$ and a death rate of $(\mu + i\theta(1))/2$ $\forall i \geq 1$. Which literally means that given the initial condition, the distribution of T (5.18) is given by the minimum of the first passage times σ_A and σ_B. Moreover, quantity 5.18 is evaluated as $\mathbf{Pr}[\sigma_A < \sigma_B]$. From Equation 5.13, we see that the conditional Laplace transform of $\sigma_A \cdot -\sigma_B$ is given $\hat{f}_{a,1}(s)\hat{f}_{b,1}(-s)$, then from Proposition 5.2.2 we can deduce that the Laplace transform of cdf of $\sigma_A \cdot -\sigma_B$ is given by Equation 5.21. The probability of interest, by definition, the value of cdf of $\sigma_A \cdot \sigma_B$ at 0, hence the value of inverse Laplace transform of quantity 5.21 evaluated at 0. □

5.4 Executing an Order before Mid-price Move

Market participants who submit a limit order stand a chance of getting a better price than those who submit market orders, however they are at risk of facing non-execution or even a phenomenon known as the *winner's curse*, where the trader tends to pay more than anticipated or perhaps the value of the asset becomes less than the bidder anticipated. In this continuous double auction, whiles market orders are executed with certainty, limit orders wait in the book in queues until they are matched with opposite order or cancelled. Hence in trading, the probability that a limit order is executed before the price move is very essential in judging the choice between a limit order and market order.

We compute the probaility that an order placed at the bid is executed before any movement in the mid-price, given that the trader is patient and does not cancel the order. We consider only the case when $S = 1$, the probability that the order is executed before the mid-price moves away from the desired price, given that the order is not cancelled. Though for now we focus on an order placed at the bid price; since our model is symmetric in bids and asks, our results also holds for orders placed at the ask price.

We let NC_b (NC_a) denote the event that an order that never gets cancelled is placed at the bid (ask) at time 0. Then the probability that an order placed at the bid is executed before the mid-price moves is given by

$$\mathbf{Pr}[\epsilon_B < T | X_b(0) = b, X_A(0) = a, s(0) = S, NC_b]. \tag{5.22}$$

Proposition 5.4.1. : *Probability of order execution before mid-price moves*
We define $\hat{f}_{a,S}(s)$ *as in (5.19) and then let*

$$\hat{g}_{j,S}(s) = \prod_{i=1}^{j} \frac{(\mu + \theta(S)(i-1))/2}{(\mu + \theta(S)(i-1))/2 + s} \tag{5.23}$$

Given that $j \geq 1$*, and* $\Lambda_S \equiv \sum_{i=1}^{S-1} \lambda(i)/2$ *Then (5.22) is given by the inverse Laplace transform of*

$$\hat{F}_{a,b,S}(s) = \frac{1}{s}\hat{g}_{b,S}(s)\left(\hat{f}_{a,S}(2\Lambda_S - s) + \frac{2\Lambda_S}{2\Lambda_S - s}(1 - \hat{f}_{a,S}(2\Lambda_S - s))\right) \tag{5.24}$$

evaluated at 0. When $S = 1$*, (5.24) reduces to*

$$\hat{F}_{a,b,1}(s) = \frac{1}{s}\hat{g}_{b,1}(s)\hat{f}_{a,1}(-s). \tag{5.25}$$

Proof. We consider the special case where $S = 1$. We construct $\tilde{X}_A$ and $\tilde{W}_B$ as described in

Lemma 5.3.1. Let $T' \equiv \min(\epsilon_B, T)$ denotes the first time when either the process $\tilde{W}_B$ empties to zero or the mid-price changes. Given that an infinitely patient trader placed an order at the bid price at time 0, then T' is the first time either the order gets executed or or the mid-price changes. From previous definition (see (5.5) and also Proposition 5.4.1) ϵ_B is given by the sum of b independent exponentially distributed random variables with parameters $(\mu+(i-1)\theta(1))/2$, $\forall i \in [1, b]$ and independent of $\tilde{X}_A$. Thus, the conditional Laplace transform of ϵ_B given the initial condition is given by Equation 5.23. Recall that

$$\epsilon_B = \sum_{i=1}^{b} W_{B_i}$$

where $W_{B_i} \sim Exp[(\mu+(i-1)\theta(S))/2] \quad \forall i \in [1, b]$ then from quantity 5.3 the Laplace transform of ϵ_B can be presented as

$$\begin{aligned}
\hat{g}_{b,S}(s) &= \mathfrak{L}[\epsilon_{B1}] \cdot \mathfrak{L}[\epsilon_{B2}] \cdots \mathfrak{L}[\epsilon_{Bb}] \\
&= \int_0^\infty \left(\frac{\mu}{2}\right) e^{-\left(\frac{\mu}{2}\right)t} e^{-st} dt \int_0^\infty \left(\frac{\mu+\theta(S)}{2}\right) e^{-\left(\frac{\mu+\theta(S)}{2}\right)t} e^{-st} dt \cdots \\
&\quad \int_0^\infty \left(\frac{\mu+(i-b)\theta(S)}{2}\right) e^{-\left(\frac{\mu+(i-b)\theta(S)}{2}\right)t} e^{-st} dt \\
&= \prod_{i=1}^{b} \frac{(\mu+\theta(S)(i-1))/2}{(\mu+\theta(S)(i-1))/2+s}.
\end{aligned}$$

We generalize the above quantity and represent it as in Equation 5.23.

The mid-price changes before ϵ_B only when $\sigma_A < \epsilon_B$, given that $S = 1$, which implies that the quantity (5.22) can be represented as $\mathbf{Pr}[\epsilon_B < \sigma_A]$. Hence using (5.3) and Proposition 5.2.2 together with the conditional Laplace transforms of ϵ_B and σ_A, shown by (5.23) and (5.19) respectively, we obtain quantity (5.25). We then compute the desired probability by computing the inverse Laplace transform of (5.25), evaluated at 0. □

Chapter 6

Numerical Results & Conclusion

The proposed stochatic moldel by Cont et al. [2010] with the Laplace transform described in Chapter 5 allow us to simulate the dynamics of the order book and also compute quantities of interest using the parameters μ, λ and θ which are estimated from the order flow. In this chapter we compute these quantities of interest, namely: probability of increase in mid-price and the probability of executing a bid order before the mid-price moves, for AstraZeneca PLC. We also compare the conditional probabilities of events in this model to corresponding empirical frequencies in the data. Finally we compare conditional probabilities of these events estimated from the Laplace transform to the simulation results by Monte Carlo.

6.1 Conditional Distributions

The key essense for simulating a high-frequency trading is to estimate conditional distributions of different events describing the order book to predict the short-term behaviour of various quantities which might be useful in trade execution and algorithmic trading. A good estimation of these conditional distributions give one the ability to predict their behaviour in the short-term [Cont et al., 2010].

6.1.1 One-step Transition Probabilities

In attempt to test our model's usefulness and reliability in predicting the short-term behaviour of the order book, we compare the one-step transition probabilities of our model to corresponding empirical frequencies, in this case we consider the probability that the number of orders at a given price level increases given that it changes.

We let T_m to be the time of the m^{th} event in the order book such that :

$$T_0 = 0 \quad \text{and} \quad T_{m+1} \equiv \inf\{t \geq T_m | X(t) \neq X(T_m)\}. \tag{6.1}$$

The probability that the number of orders at a distance i from the opposite best quotes moves from n to $n+1$ at the next change is given by

$$\begin{aligned} P_i(n) &\equiv \mathbf{Pr}[Q_i^A(T_{m+1}) = n+1 | Q_i^A(T_m) = n, Q_i^A(T_{m+1}) \neq n, s(T_m) = 1] & (6.2) \\ &= \begin{cases} \frac{\lambda(1)/2}{\lambda(1)/2+\mu/2+n\theta(1)/2} & \text{where} \quad i = 1 \\ \frac{\lambda(i)/2}{\lambda(i)/2+n\theta(i)/2} & \text{where} \quad i > 1 \end{cases} & (6.3) \end{aligned}$$

For instance, we consider the situation where $i = 1$, The next change in Q_1^A is an increase if there is an arrival of a limit ask order at price level where Q_1^A occured before the occurence of any other event, be it a market order or a cancellation. However since the arrival of a limit order at the price level of Q_1^A occurs at a rate of $\lambda(1)/2$ and a cancellation or market order occur at rate of $(\mu + n\theta(1))/2$, it follows from assumption that the probability of a limit ask order occuring first is

$$\frac{\lambda(1)/2}{(\lambda(1) + \mu + n\theta(1))/2}.$$

$\hat{Q}_i^B$ ($\hat{Q}_i^A$) denotes the empirically observed number of bid (ask) orders at a distance i from the opposite best quote at time t. An estimator for the above probability (quantity 6.2) is given by

$$\hat{P}_i(n) \equiv \frac{\hat{B}_{up} + \hat{A}_{up}}{\hat{B}_{change} + \hat{A}_{change}} \tag{6.4}$$

where

$$\begin{aligned} \hat{B}_{up} &= \left| \{m | \hat{Q}_i^B(\hat{T}_m) = n, \hat{Q}_i^B(\hat{T}_{m+1}) > n\} \right| \\ \hat{A}_{up} &= \left| \{m | \hat{Q}_i^A(\hat{T}_m) = n, \hat{Q}_i^A(\hat{T}_{m+1}) > n\} \right| \\ \hat{B}_{change} &= \left| \{m | \hat{Q}_i^B(\hat{T}_m) = n, \hat{Q}_i^B(\hat{T}_{m+1}) \neq n\} \right| \\ \hat{A}_{change} &= \left| \{m | \hat{Q}_i^A(\hat{T}_m) = n, \hat{Q}_i^A(\hat{T}_{m+1}) \neq n\} \right| \end{aligned}$$

Figure 6.1 displays $P_i(n)$ and $\hat{P}_i(n)$ $\forall i \in [1, 10]$ for AstraZeneca PLC. We could see that these probabilities are reasonably close in most cases, indicating that the short term dynamics of the order book is described by our model.

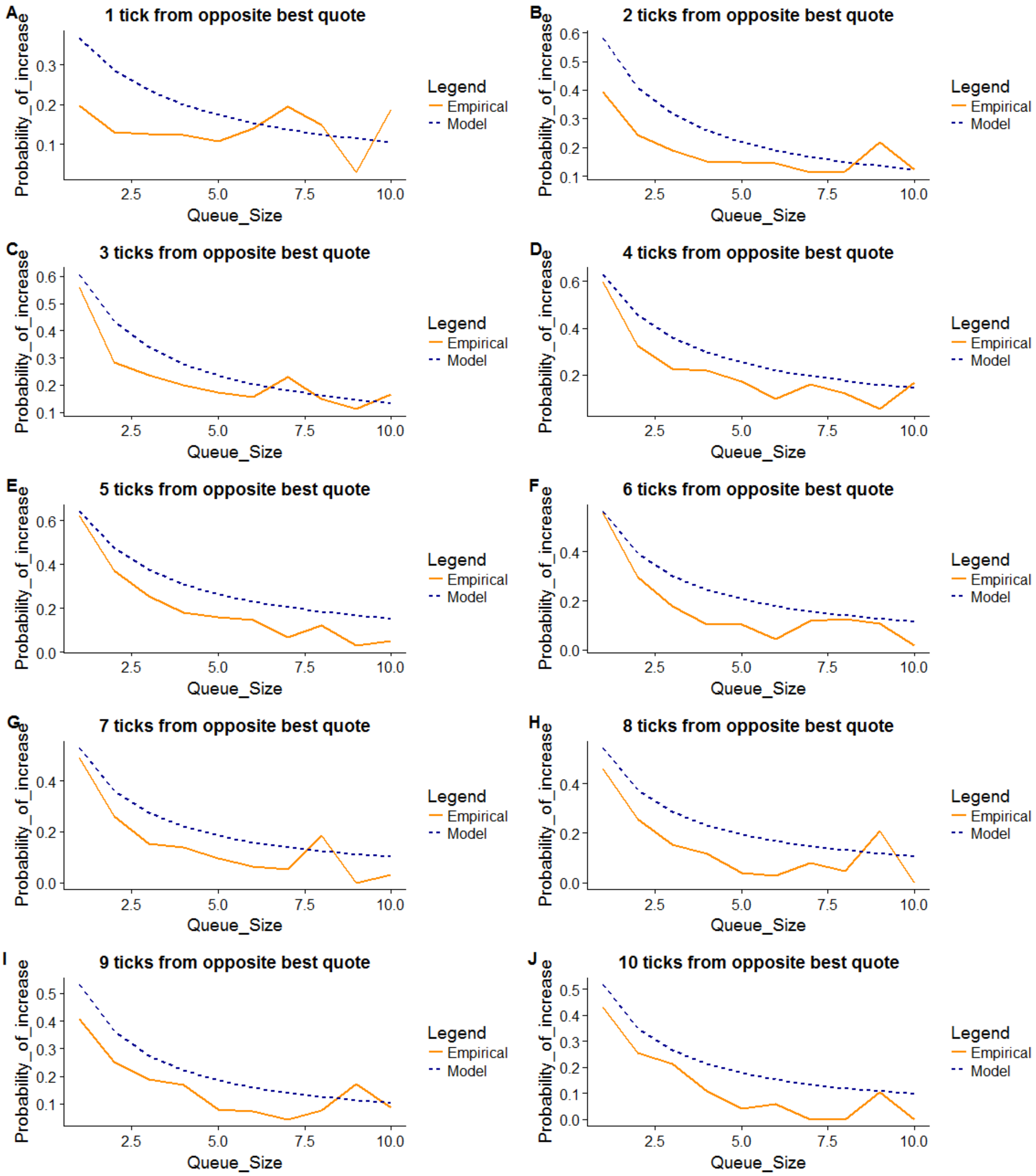

Figure 6.1: Probabilities of increase in the number of orders at distance i from the opposite best quote in the next change, $\forall i \in [1, 10]$.

6.1.2 Direction of Price Movement

Here and in the next section, we look at the computation of the conditional probabilities using the inversion of Laplace transform which were computed in **R**. For probabilities of increase in the mid-price, we realize that $\forall a = b$ the probability is always 0.5 and these probabilities behave well when $a \neq b$. We computed these probabilities of mid-price increase and bid execution using a Monte Carlo simulation of 30 000 replications and Laplace transform as described in Chapter 5.

Tables 6.1 and 6.2 show that the theoretical results and our implementation of the inverse Laplace transform are correct.

	a									
b	1	2	3	4	5	6	7	8	9	10
1	0.50	0.68	0.77	0.82	0.85	0.88	0.89	0.91	0.92	0.92
2	0.32	0.50	0.61	0.68	0.73	0.77	0.80	0.82	0.84	0.85
3	0.23	0.39	0.50	0.58	0.64	0.68	0.72	0.75	0.77	0.79
4	0.18	0.32	0.42	0.50	0.56	0.61	0.65	0.68	0.71	0.73
5	0.15	0.27	0.36	0.44	0.50	0.55	0.59	0.63	0.66	0.68
6	0.12	0.23	0.32	0.39	0.45	0.50	0.54	0.58	0.61	0.64
7	0.11	0.20	0.28	0.35	0.41	0.46	0.50	0.54	0.57	0.60
8	0.09	0.18	0.25	0.32	0.37	0.42	0.46	0.50	0.53	0.56
9	0.08	0.16	0.23	0.29	0.34	0.39	0.43	0.47	0.50	0.53
10	0.08	0.15	0.21	0.27	0.32	0.36	0.40	0.44	0.47	0.50

Table 6.1: Probabilities of an increase in mid-price: Laplace transform results.

	a									
b	1	2	3	4	5	6	7	8	9	10
1	0.50	0.68	0.77	0.82	0.85	0.88	0.89	0.91	0.92	0.92
2	0.31	0.50	0.61	0.68	0.73	0.77	0.80	0.82	0.84	0.85
3	0.23	0.39	0.50	0.58	0.64	0.68	0.72	0.74	0.77	0.79
4	0.18	0.32	0.42	0.50	0.56	0.61	0.64	0.68	0.71	0.74
5	0.15	0.27	0.36	0.44	0.50	0.55	0.59	0.63	0.65	0.68
6	0.12	0.24	0.31	0.39	0.44	0.50	0.54	0.57	0.61	0.64
7	0.11	0.20	0.28	0.35	0.41	0.46	0.50	0.53	0.58	0.60
8	0.09	0.18	0.25	0.32	0.38	0.42	0.47	0.50	0.53	0.56
9	0.08	0.16	0.23	0.29	0.34	0.39	0.44	0.46	0.51	0.53
10	0.07	0.15	0.21	0.27	0.32	0.37	0.40	0.44	0.47	0.50

Table 6.2: Probabilities of an increase in mid-price: Simulation results.

6.1.3 Executing an Order before Mid-price Change

we compute and compare the conditional probabilities of executing a given bid before any increase in the mid-price, for different values of a and b given that $S = 1$. Tables 6.3 and 6.4

show the results, again one can see how close the simulation results are to the Laplace transform result.

b \ a	1	2	3	4	5	6	7	8	9	10
1	0.43	0.27	0.19	0.15	0.12	0.10	0.09	0.08	0.07	0.07
2	0.62	0.44	0.34	0.27	0.23	0.20	0.17	0.15	0.14	0.13
3	0.72	0.55	0.44	0.37	0.32	0.28	0.25	0.22	0.20	0.18
4	0.78	0.63	0.53	0.45	0.39	0.35	0.31	0.28	0.26	0.24
5	0.83	0.69	0.59	0.51	0.45	0.40	0.37	0.33	0.31	0.28
6	0.86	0.74	0.64	0.57	0.50	0.46	0.41	0.38	0.35	0.33
7	0.88	0.77	0.68	0.61	0.55	0.50	0.46	0.42	0.39	0.36
8	0.89	0.80	0.71	0.64	0.59	0.54	0.50	0.46	0.43	0.40
9	0.91	0.82	0.74	0.68	0.62	0.57	0.53	0.49	0.46	0.43
10	0.92	0.84	0.76	0.70	0.65	0.60	0.56	0.52	0.49	0.46

Table 6.3: Probabilities of executing a bid order before a change in mid-price: Laplace transform results.

b \ a	1	2	3	4	5	6	7	8	9	10
1	0.43	0.26	0.20	0.14	0.12	0.10	0.09	0.08	0.07	0.07
2	0.62	0.43	0.33	0.28	0.23	0.20	0.18	0.15	0.14	0.13
3	0.72	0.56	0.44	0.37	0.32	0.28	0.24	0.22	0.20	0.18
4	0.79	0.64	0.53	0.45	0.39	0.34	0.31	0.28	0.25	0.24
5	0.82	0.69	0.60	0.51	0.45	0.40	0.36	0.33	0.31	0.28
6	0.85	0.74	0.64	0.57	0.51	0.45	0.41	0.38	0.35	0.33
7	0.88	0.77	0.68	0.61	0.55	0.50	0.46	0.43	0.39	0.36
8	0.89	0.80	0.72	0.65	0.58	0.54	0.49	0.46	0.43	0.40
9	0.91	0.82	0.74	0.68	0.62	0.58	0.53	0.49	0.46	0.43
10	0.92	0.84	0.77	0.71	0.65	0.60	0.56	0.52	0.49	0.46

Table 6.4: Probabilities of executing a bid order before a change in mid-price: Simulation results.

6.2 Conclusion

We have studied the dynamics of the order book for AstraZeneca PLC by employing the stylized stochastic model proposed by Cont et al. [2010]. In the stylized stochastic model different types of event –market orders, limit orders and cancellations– are described by an independent Poisson processes which is analogous to a queuing system.

The model is based on observable quantities and the parameters can be estimated from observation of events in the order book. It allows the analytical computation of different conditional probabilities of the order book events using the Laplace transform method. This model adequately captures the short-term behaviour of the order book.

From our numerical results we could see that the conditional distribution of various quantities of interest agree with the corresponding empirical distributions for parameters estimated from our data.

The ability of our model to compute realiable conditional distribution is very useful for short-term prediction and also electronic trading strategies. This study also shows how far the stochastic model can go in reproducing dynamics of the limit order book bypassing behavioural assumptions of market participants.

This model is not always compactible with some order book data. For instance in our first attempt to model with a 2006 order book data from Vodafone, it was realized that most of the outstanding orders at given price levels were not emptying to zero due to inconsistencies in the proportionality of cancellation. In our case at the distance 1 from the best bid and ask we had a limit order rate of 7.38 and a cancellation rate of 0.23. We see how how small the cancellation rate is relative to the limit order rate. Hence with a constant market rate of 2.39 most of the outstanding orders could not empty to zero before the close of day.

Taking all into account, it is certainly worth considering using the model for developing trade algorithms and study in more detail, how the model can be improived for stocks like Vodafone, where the current model does not give reasonable results.

Bibliography

J. Abate and W. Whitt. Numerical inversion of laplace transforms of probability distributions. *ORSA Journal on Computing*, vol.7(1):36–43, 1995.

J. Abate and W. Whitt. Computing laplace transforms for numerical inversion via continued fractions. *INFORMS Journal on Computing*, 4(11):394–405, 1999.

Jean-Philippe Bouchaud, Marc Mézard, and Marc Potters. Statistical properties of stock order books: empirical results and models. *Quantitative Finance*, 2:251–256, 2002.

K.J. Cohen, R.M. Conroy, and S.F. Maier. Order flow and the quality of the market. *In : Y. Amihud, T. Ho, R. Schwartz, Market Making and the Changing structure of the Securities Industry, Lexington, MA: Lexington-Heath Books Co*, pages 93–112, 1985.

S.N. Cohen and L. Szpruch. A limit order book model for latency arbitrage. *Mathematics and Financial Economics*, 6:211–227, 2012.

R. Cont, S. Stoikov, and R. Talreja. A stochastic model for order book dynamics. *Operations Research*, vol.58(no.3):217–224, 2010.

I Domowitz and J. Wang. Auctions as algorithms: Computerized trade execution and price discovery. *Journal of Economic Dynamics and Control*, vol.18(no.1):29–60, 1994.

Thierry Foucault. Order flow composition and trading costs in a dynamic limit order market. *Journal of Financial Markets*, vol.2(no.2):99– 134, 1999.

L. Glosten. Is the electronic open limit order book inevitable? *Journal of Finance*, 49:1127–1161, 1994.

G. Gu and Wei-Xing Zhou. Emergence of long memory in stock volatility from a modified mike-farmer model. *EPL (Europhysics Letters)*, vol.86:48002, 2009.

B. Hollifield, R. A. Miller, and P Sandas. Empirical analysis of limit order markets. *The Review of Economic Studies Limited*, vol.71(0034-6527/04/00411027):1027–1063, 2004.

G. Jiang, S. Wang, and H. Dong. A survey of limit order book modeling in continuous auction market. *National Natural Science Foundation for Young Scholars of China*, (No.11001179), 2011.

Sham M. Kakade, Michael Kearns, Yishay Mansour, and Luis E. Ortiz. Competitive algorithms for vwap and limit order trading. *Social and Behavioral Sciences Economics*, 2004.

A.N. Kercheval and Y. Zhang. Modelling high-frequency limit order book dynamics with support vector machines. *Quantitative Finance*, vol.15(No.8):1315–1329, 2015.

Andrei Kirilenko, Albert S. Kyle, Mehrdad Samadi, and Tugkan Tuzun. The flash crash: The impact of high frequency trading on an electronic market. *Working Paper*, 2011.

R. Kozhan and M. Salmon. The information content of a limit order book: The case of an fx market. *Journal of Financial Markets*, Vol.15(1):1–28, 2012.

D. Liu. Particle-scale modelling of financial price dynamics. *Communications in Nonlinear Science and Numerical Simulation.*, (43):282–295, 2017.

Hugh Luckock. A steady-state model of the continuous double auction. *Quantitative Finance*, Vol.3:385–404, 2003.

S. Mike and J.D. Farmer. An empirical behavioral model of liquidity and volatility. *Journal of Economic Dynamics and Control*, vol.32:200–234, 2007.

Christine A. Parlour. Price dynamics in limit order markets. *The Society for Financial Studies*, vol.11(no.4):pp. 789–816, 1998.

K. Rock. The specialist's order book and price anomalies. *Working paper, Harvard University*, 1996.

I. Rosu. A dynamic model of the limit order book. *Published by Oxford University Press on behalf of The Society for Financial Studies.*, vol.22(no.11):pp. 4601–4641, 2009., 2009.

D Seppi. Liquidity provision with limit orders and a strategic specialist. *Review of Financial Studies*, 10:103–150, 1997.

E. Smith, J.D. Farmer, L. Gillemot, and Krishnamurthy. Statistical theory of the continuous double auction. *Quantitative Finance*, vol.3:481–514, 2003.

I.M. Toke. The order book as a queueing system: average depth and influence of the size of limit orders. *Quantitative Finance*, Vol.15(No.5):795–808, 2015.

Antti Tuominen. The role of high frequency trading in limit order book activity: Evidence from helsinki stock exchange. *Aalto University School of Economics*, 2012.

Non-exclusive licence to reproduce thesis and make thesis public

I Dominic Isaac Andoh (date of birth: 08.11.1989),

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

   (a) reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and

   (b) make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,"Limit Order Book Modelling –A Stochastic Approach–", supervised by Assoc. Prof. Raul Kangro,

2. I am aware of the fact that the author retains these rights.

3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu,16.05.17