ESTIMATING LIQUIDITY RISK FROM HIGH FREQUENCY TICK DATA

A dissertation submitted to the Centre for Computational Finance and Economic Agents (CCFEA), University of Essex

In partial fulfillment of the requirements for the degree of MSc. Computational Finance

Supervisor: Dr. Steve Phelps

By

Bruce I. Uponi
August, 2012
ACKNOWLEDGEMENT

This research project would not have been possible without the support of many people. First and foremost, I wish to express my profound gratitude to my supervisor, Dr. Steve Phelps who was more than abundantly helpful and offered invaluable assistance, support and guidance all through the project. A special gratitude also to Neil Rayner whose very kind assistance was enough to fix the bugs in the SETS Order Book Rebuilder Application. Special thanks also to Dr. Win Ng Lon for granting me a session that caused a desired paradigm shift in the project by critically examining my data. I would also not forget to specially thank Prof. Edward Tsang who helped to inspire this topic via a skype call conversation. I cannot thank you all enough.

Very much worth mentioning, is my heartfelt gratitude to my entire family, my dad especially who has been my support all through my academic journey. Thank you dad for being there. Lastly, I owe this to the source of my strength, the one with all wisdom and power, my maker, for the grace of completion which He has bestowed me. Thank you Lord.
ABSTRACT

In this dissertation I build on existing work that attempts to provide quantitative metrics for liquidity risk by analyzing supply and demand in the market. Existing approaches are based on macro-economic models which make many assumptions. However, in an order-driven market such as the London Stock Exchange (LSE), the limit order book contains precise data on the marginal supply and demand expressed in terms of bids and asks offers from both buyers and sellers. Using high-frequency data from the London Stock exchange, I estimate the liquidity risk in the market by reconstructing the limit-order book and the corresponding marginal supply and demand curves. This is a first step towards developing high-frequency trading strategies that incorporate liquidity-risk metrics.
# TABLE OF CONTENT

1. **INTRODUCTION**  
   1.1. Properties of Liquidity  
   1.2. Price Formation Process  
   1.3. Summary  

2. **LITERATURE REVIEW**  
   2.1. Spread Measures  
   2.2. Price Impact Measures  
   2.3. Volume based Measures  
   2.4. Dynamic Semi-parametric Factor Model (DSFM)  
   2.5. VaR Measures  
   2.6. Summary  

3. **METHODOLOGY**  
   3.1. Order Book Reconstruction  
   3.2. The New Liquidity Formalism  
   3.3. Marginal Supply Demand Curve (MSDC)  
   3.4. The New Portfolio Framework  

4. **DATA**  

5. **ANALYSIS & EMPIRICAL RESULTS**  
   5.1. MSDCs of Stocks (AZN & BATS)  
   5.2. Liquidity Risk Estimation  
   5.3. Summary  

6. **CONCLUSION**  

7. **REFERENCES**  

8. **APPENDIX**
LIST OF TABLES

1. Table 3.1: Outline of the SETS Simulator Order Book reconstruction process 18
2. Table 3.2: List of prices and MSDC for a stock. 21
3. Order book snapshots for AZN and BATS stocks at 02/06/2008 08:30 38

LIST OF FIGURES

1. Figure 1.1: Aspects of liquidity in a static state of the limit order book. 4
2. Figure 2.1: VaR with Spread Adjustment. 14
3. Figure 4.1: LSE SETS market operation. 25
4. Figure 4.2: Order Book Data generation process. 26
5. Marginal Supply and Demand Curves for AZN and BATS stocks. 40
6. Short 10 various quantities of AZN stock at 12:30 and 16:30. 42
7. Long AZN stock and Short BATS – 10 possible quantities of each at 12:30 and 16:30. 43
8. Long AZN stock and Short BATS – 10x10 possible quantities of each at 12:30 and 16:30 44
9. Liquidation MtM U (p) function for MSDC showing a concave pattern 45
10. Liquidation MtM L (p) function for MSDC showing a concave pattern 45
1. INTRODUCTION

Liquidity, until recently assumed less priority in the checklist of risk concerns of investors and the financial market in general. The fatal investment experience of 2007 and 2008 has taught all financial market stakeholders a critical lesson. Now, the "focal point" of almost every investment from institutional investors to private investors and traders inclusive is Liquidity – the risk of endogenous and/or exogenous factors that affect the flow of cash and consequently cause market frictions or slow down. Liquidity is important, and it needs to be identified and managed like any other investment risk. The Lehman episode especially, triggered the pressing need for formal liquidity requirements. Consequently, the Basel Committee has made liquidity risk a major focus of Basel III.

While liquidity still substantially remains an elusive concept with various definitions across varying micro and macro market structures, we confine our definition to the ability to trade large quantities quickly, at a low cost, and without moving the price of the asset adversely. Also, a market characterized by the ability to buy and sell with ease clearly depicts a liquid market. But because all markets are not liquid all of the time, we can confidently make a case for illiquid markets. It will be an economic fallacy to assume that markets are constantly liquid. Acerbi & Scandolo [5][7] argues that the Portfolio Theory is based on this assumption – that markets are perfectly liquid, and therefore betrays hidden assumptions of perfect market liquidity by not sufficiently providing a basis to consider generally illiquid markets. Hence, the Portfolio Theory has to be re-written. Having established the existence of market illiquidity, the idea of liquidity
risk comes to bear. A generalized notion of liquidity risk is broadly categorized into the following three contexts as described by Acerbi & Scandolo [5]:

1. Portfolio Liquidity Risk – the risk that a portfolio runs out of cash “money” necessary for future payments (*the “treasurers’ point of view”*).

2. Market Liquidity Risk – the risk of buying or selling on a market which is shallow compared to the trade size (*the “traders’ point of view”*).

3. Systemic Liquidity Risk – the risk that the liquidity circulating in the economy is dried up (*the “policy-maker point of view”*). This is a gradual resultant of (1) and (2).

Acerbi & Scandolo [5][7] upholds that liquidity risk is a combination of all three aspects without any exclusion and hence proposes a formalism satisfying these three contexts of liquidity risk as a basis for evaluating liquidity risk problems. Using this formalism, I focus primarily on market liquidity risk in high-frequency finance with the limit order book as a case study, and seek to estimate the latent liquidity risk in the event of liquidating a hypothetical trader’s portfolio. Prior to this, I discuss some selected models and approaches to liquidity risk measurement in the limit order book by other authors.

The primary goal of this work however, is to analyze and estimate liquidity risk in High Frequency Financial market where information regarding market activities is very much obtainable from the limit order book. With regards to several approaches and models that exists on modeling and evaluating liquidity and its attendant risks, more emphasis is put on reconstructing the limit order book and deriving the corresponding Marginal Supply and Demand Curves (MSDCs) from the reconstructed order book by capturing order book data at time specific states of the order book which directly imply the snapshots of orders placed in the
order book ex-ante. Going further, the cost of liquidating a portfolio is estimated using the formalism put forward by Acerbi & Scandolo [5][7] by evaluating the lower and upper Mark-to-market of a portfolio of two assets.

### 1.1 Properties of Liquidity

There are four aspects of liquidity which are reflected in an order book. Kyle [22], Hasbrouck & Schwartz [16], and Roll [29] were able to reveal these aspects as the following:

1. **Tightness** – the difference between Bid and Ask prices. It is measured in terms of the bid-ask spread. In a perfectly liquid market with respect to tightness, the spread would be zero and traders can buy or sell at the same price.

2. **Market depth** – the ability to buy or to sell a certain amount of an asset without influencing the quoted market price. The key focus here is the market volume. A shallow market depth is a sign of illiquidity thus; an infinite depth at the quoted price is expected of a perfectly liquid market.

3. **Resilience** – this is the mean-reverting expectation of the market price following a random shock in the transaction flow. If there is no adverse impact on price, price would revert to the former state it was before the impact sustained from previous transaction i.e. equilibrium. Thus in highly liquid markets, with respect to resiliency, prices absolutely bounce back to their equilibrium level.
4. **Immediacy of Execution**\(^1\) – the time between sending the order and its execution at the prevailing, instant price. Large limit orders visible to all market participants are prone to destabilize market equilibrium by impacting prices even without being executed.

Figure 1.1 shows these liquidity aspects in a static state of the limit order book. On the vertical axis is the price, the Bid and Ask prices respectively. The difference in prices is the bid-ask spread which is a measure of the tightness in price. On the horizontal axis is the size or volume. The sum of the Bid and Ask volumes is a measure of the market depth, and the diagonal is the market resilience i.e. the mean reversion of prices to their equilibrium level, which also reveals the price elasticity of supply and demand. This aspect is further elaborated in the Marginal Supply and Demand Curve (MSDC) in Figure 3.1.

\[\text{Figure 1.1: Aspects of liquidity in a static state of the limit order book. (Source: Bervas, 2006)}\]

\(^1\) Kyle’s (1985) work does not include Immediacy of Execution.
1.2 Price Formation Process

The market structure and the trading mechanisms of the particular market in consideration are the underlying factors fundamentally responsible for its price formation process. To analyze liquidity in order driven markets, the price formation process is essentially critical in understanding how liquidity risk crystallizes from the systemic interactions of all market elements. Markets must have assets, assets must have prices, prices are decided by traders, and traders decide prices based on market stimulus. Thus, the key elements of a market include assets, prices and traders. These three elements modulate and evolve the price formation of any market, order driven markets inclusive. In order driven markets, participants typically do not trade directly with each other but quote their buy or sell orders on display for all participants in the market via the order book. The availability of all quotes to all market participants increases the chance of participants to find the best prices for their orders. The trading system in an order driven market such as the exchange facilitates the determination of the price in real time. This mechanism is called an auction. Through auctions, market participants quote prices to sell or buy an asset. The combination of all sale and purchase orders at a given time can be represented by a list of ask prices and bid prices with corresponding trading volumes. An ask price is the price that a seller is willing to accept for an asset while a bid price is the price that a buyer is willing to pay for an asset. The difference between the lowest ask price and the highest bid price will be called the bid-ask spread. The best ask and bid prices are at the top of the order book waiting to be matched, and then a trade is done. Exchanges have automatic order matching algorithms embedded in their order book system that makes this possible.
1.3 Summary

From the order book we can infer the bid-ask spread as a proxy for liquidity. Although some authors\(^2\) have argued that this is a rather trivial proxy for liquidity without considering the market depth, as it does not reflect the liquidity supply beyond the best prices. This will be discussed in detail in the next chapter. By the bid-ask spread, a typical liquidity risk present in trading the order book can be illustrated with a simple case where there are two ask orders in the order book – 1000 shares of IBM at $100.50 and a second best ask quote of 1500 shares of IBM at $100.55. If a trader was to place an order to buy 2000 shares of IBM, the first 1000 shares would be cleared at $100.50 and the 1000 shares remaining would be served with the next best ask price – at $100.55. The effective price at purchase is hence $100.53. This requires the trader to make prior liquidity provision to be able to fund his position for whatever purpose he intended for the purchase at the volume requested. Liquidity issues of this manner do occur and the effect is significant when large order volumes are transacted. The cost of liquidity is hence the ultimate question. Subsequent chapters will be dedicated to discuss this in detail. Chapter 2 will discuss the approaches taken by some authors in measuring liquidity risk. Chapter 3 presents the model adopted to estimate liquidity risk in my work. Data and Results are discussed in Chapters 4 and 5 respectively, and Chapter 6 concludes.

2. LITERATURE REVIEW

After a critical elicitation of liquidity and liquidity risk in the introductory chapter, the actual task of measuring liquidity risk can now be adequately confronted. I begin by reviewing existing quantitative methods and techniques pertaining to liquidity risk and its measurement. With so many models existing in this regard, a relevant selection of them are discussed herein. The selected models include – simple spread measures, price impact measures, volume based measures, Dynamic Semi-parametric Factor Model (DSFM), and finally the VaR based approach. The correctness and efficiency of these models were not tested due to time constraints, hence are not captured in my report. They are presented here to show the ideological distinction between their approach and the adopted approach used in my research in chapter 3.

2.1 Spread measures

The bid-ask spread is a natural measure of liquidity commonly known as a proxy for liquidity. With the following notations, the spread measure follows:

\[ s_i := a_i - b_i , \]  

(2.1.0)

\( a_i \) and \( b_i \) denote the ask and bid prices at time \( t_i \). Also, the use of modified bid-ask spread as seen from Goyenko et al [] is the result of the effective spread given by

\[ s_{e,i} := 2|\ln p_i - \ln mq_i| , \]  

(2.1.1)

where \( mq_i = 0.5(a_i - b_i) \) denotes the mid-price. The effective spread measures the “absolute” distance between the transaction price and the mid-price and thus the effective transaction costs
implied by a trade. If the effective spread is aggregated over time, we can compute a stock’s
dollar-volume-weighted average trading costs. By this we can derive the realized spread given by

$$s_{r,i} := 2|\ln p_i - \ln mq_{i+jx}|,$$  \hspace{1cm} (2.1.2)

where $j_x$ is the next trade occurring at the $x^{th}$ minute interval, and $x$ indicates the sampling
frequency used in computing usually 5minutes. This measure gives the spread which is realized
after $x$ minutes and thus reflects the transaction costs which apply if the stock would be bought
or sold back after this time. If an identification of the trade direction is possible, the realized
spread measure can be modified as

$$s_{r,i} := \begin{cases} 2(\ln p_i - \ln mq_{i+jx}) & \text{if } ith \text{ trade is a Buy} \\ 2(\ln mq_{i+jx} - \ln p_i) & \text{if } ith \text{ trade is a Sell} \end{cases}$$ \hspace{1cm} (2.1.3)

A simple measure of the permanent price impact of a trade in $i$ is then given by the difference
between the effective spread and the realized spread, i.e.

$$s_{e,i} - s_{r,i} = 2|\ln mq_{i+jx} - \ln mq_i|.$$  \hspace{1cm} (2.1.4)

\section*{2.2 Price impact measures}

A simple model of price dynamics proposed by Roll [] takes transaction cost into account. It
assumes that mid prices follow a random process

$$mq_i = mq_{i-1} + \varepsilon_i,$$  \hspace{1cm} (2.2.0)

where $\varepsilon_i$ is the white noise process with prices given by

$$p_i = mq_{i-1} + cy^b_i,$$  \hspace{1cm} (2.2.1)

where
\[ y_i^b = \begin{cases} 1 & \text{if } i \text{ is a Buy} \\ -1 & \text{if } i \text{ is a Sell} \end{cases}, \]  

(2.2.2)

and \( c \) reflects the effective trading costs corresponding to the half spread. The model then implies

\[ \Delta p_i = mq_i + cy_i^b - (mq_{i-1} + cy_{i-1}^b) = c\Delta y_i^b + \varepsilon_i. \]  

(2.2.3)

and \( c = \sqrt{-\text{Cov}[\Delta p_i - \Delta p_{i-1}]} \). The effective transaction cost can then be measured as the square root of the negative first order autocovariance of transaction price changes. This measure is only valid as long as the first order autocovariance is negative. Generalizations of the Roll model are proposed by Hasbrouck [19] and later extended in Hasbrouck [20] to estimate transaction costs using daily data with the use of a Gibbs sampler. Similarly, Amihud [33] proposes an illiquidity measure given by

\[ I = \frac{|r_d|}{v_d}, \]  

where \( d \) indexes trading days, \( r_d \) denotes daily log returns and \( v_d \) is the daily trading volume. Accordingly, the Amihud measure gives the average price change per trade size and thus reflects the average price impact of a transaction unit. Correspondingly,

\[ L = I^{-1} = \frac{v_d}{|r_d|}, \]  

(2.2.5)

is a liquidity measure and is sometimes referred to as Amivest liquidity ratio (coined after the management firm that developed it). However, the price impact measures discussed so far are “low-frequency” measures in the sense that they are computed based on aggregated data. Correspondingly, the high frequency measures are traditionally obtained by price impact coefficients in regressions of trade-to-trade price changes on (signed) trading volume. If \( m_i \) is
denoted as the expected value of the stock, conditional on the information set at \( i \). \( m_i \) assumedly evolves as

\[
mq_i = mq_{i-1} + \epsilon_i ,
\]

(2.2.6)
as suggested by Glosten and Harris [14]; \( q_i \) denotes the signed trading volume at \( i \) and \( \epsilon_i \) is a white noise process representing (non-predictable) public signals. The coefficient \( \lambda \) represents a market depth parameter reflecting the underlying demand or supply schedule. Then, if the price process follows (2.17), price changes are given by

\[
\Delta p_i = \lambda q_i + c(y_i^b - cy_{i-1}^b) + \epsilon_i ,
\]

(2.2.7)
As suggested by Brennan et al. [], this equation can be modified to allow for different price responses to buys and sells,

\[
\Delta p_i = \alpha + \lambda_{buy}(q_i|q_i > 0) + \lambda_{sell}(q_i|q_i < 0) + c(y_i^b - cy_{i-1}^b) + \epsilon_i ,
\]

(2.2.8)
Then, \( \lambda_{buy} \) and \( \lambda_{sell} \) are referred to as “buy lambdas” and “sell lambdas” respectively, which reflect the trade-specific price impact and are easily estimated using the Ordinary Least Square (OLS) regression. A more reduced-form approach to measure the price impact of a trade is to run the regression,

\[
r_i = \lambda q_{i-1} + \epsilon_i ,
\]

(2.2.9)
and to estimate \( \lambda \) as a simple measure of the price impact. There are proposed modifications of these price impact regressions by some other authors, e.g. Hasbrouck [20]. However, all these measures require an identification of buyer and seller initiated trade. Hence, a noisy buy-sell identification can have a severe impact on their reliability.
2.3 Volume Based measures

Volume based liquidity measures are often calculated as a certain volume, or quantity of shares, at a time interval. More often they are used to estimate liquidity with respect to depth, but also incorporate time in the calculation since a higher volume in the market leads to a shorter time needed for trading a predefined amount of shares. If the volume based liquidity measures are high, it signals a sign of high liquidity. According to the Amihud measure in (2.2.4), relating the net trading volume to the corresponding price change over a fixed interval of time can be used to estimate the realized market depth. However, Engle and Lange [13] observes that using a very small time interval can lead to measurement problems due to the excess demand or the corresponding price change which can often be zero. In contrast, using longer time intervals reduces the ability of the measure to capture short-run dynamics that are of particular interest when the market is very active. For this reason, Engle and Lange [13] propose the VNET measure, which measures the log net directional (buy or sell) volume over a price duration. Using a price duration as an underlying time interval avoids the aforementioned problems and links market volatility to market depth. The VNET is computed as

\[
V_{NET_i} := \ln \left( \sum_{j=N(t_{i-1}^{dp})}^{N(t_i^{dp})} y_j^b v_j \right) \quad (2.3.0)
\]

Where \(dp\) denote price changes in \(\{t_i^{dp}\}\) which is a sequence of points and \(N(t)\) denotes the counting function associated with the transaction process. Hence, VNET measures the excess volume that can be traded before prices exceed a given threshold and therefore can be interpreted as the intensity in which excess demand flows into the market. An application of this concept in analyzing the depth of the market is found in Engle and Lange [13]. Alternative
liquidity measures that employ a volume based approach such as the Impact Cost\(^3\) (IC), equally based on trading volumes reflecting realized liquidity demand, combines both price and quantity information. It is ex ante measure of liquidity at a given quantity. The impact cost can be calculated as the percentage difference between the weighted average execution price and the pre-trade midpoint. The IC is given by

\[
IC = \frac{\sum p_i q_i - P_m Q}{P_m Q}
\]  

(2.3.1)

where \(\sum q_i = Q\), having \(q_i\) stocks at price \(p_i\) and \(P_m\) is the midpoint of the bid-ask spread at the time of the trade. Irvine, Benston & Kandel \[17\] define the cost of round trip as the sum of impact cost of ask side and impact cost of bid side.

\[
\text{Cost of Round Trip} = \text{Impact cost of Ask side} + \text{Impact cost of Bid side}
\]  

(2.3.2)

### 2.4 Dynamic Semi-parametric Factor Model (DSFM)

Härdle et al. \[15\] propose modeling the order book curves using a dynamic semi-parametric factor model (DSFM) as introduced by Park et al. \[25\]. The starting point is to assume that \(v_i^j\) follows an orthogonal \(L\)-factor model with \(L \ll J\),

\[
v_i^j = m_0 + \beta_{1i} m_1^j + \cdots + \beta_{Li} m_L^j + \varepsilon_i^j
\]  

(2.4.0)

where \(m_0\) is a constant, \(m_i^j\) is the \(j\)-level specific realization of a (time-invariant) factor loading with \(m_i = \mathbb{R}^j \to \mathbb{R}\), \(\beta_{Li}\) denotes the value of the corresponding factor at \(i\) and \(\varepsilon_i^j\) represents a white noise error term. Define \(\mathbf{m} := (m_0, m_1, ..., m_L)'\) and \(\mathbf{f}_i := (\beta_0, \beta_1, ..., \beta_{Li})'\) with \(\beta_0i = 1\).

\(^3\)S. Kumar and A. Shah (2006) gives further detail on the Impact Cost (IC) measure, with references to research done by other authors in this regard.
The DSFM is a generalization of the factor model (2.4.1) and allows the factor loadings $m_l$ to depend on explanatory variables, $z^l_i$. Accordingly, it is given by

$$v^l_i = \sum_{l=0}^{L} \beta_{il} m_i(z^l_i) + \epsilon_i^l = f_i m(z^l_i) + \epsilon_i^l \tag{2.4.1}$$

where $z^l_i, \epsilon_i^l$ and $f_i$ are assumed to be independent. More detail explanation about the DSFM is sufficiently documented in Härdle et al. [15] and Park et al. [25]. Härdle et al. [15] used this approach to model the limit order book curves of four stocks at the Australian Stock Exchange based on a 5 minutes frequency.

## 2.5 VaR measures

VaR measures of liquidity are mostly adopted using VaR as an underlying approach to consolidate the implementation of any of the fundamentally derived proxies of liquidity – spread, volume, price impact and/or time. VaR alone is not a sufficient measure of liquidity due to the negligence of the bid ask spread in its calculation Loebnitz[21], Bangia et al. [3] explains this in detail. Also, because VaR is not a coherent measure of risk, Acerbi and Scandolo [7] argues that it cannot sufficiently model liquidity risk. The Spread Adjustment approach is one out of several VaR liquidity measures existing in literature4.

Bangia et al. proposes the spread adjustment approach as a more sophisticated measure to mitigate the modeling issues of conventional VaR. They approach the problem by separating the price risk from the market “liquidity” risk. Conventional VaR calculations are used for the price

---

4 Loebnitz [21] in the paper, “Market Liquidity Risk – Elusive No More” extensively discussed the varying approaches used by several other authors to measure liquidity risk by incorporating VaR.
risk while an adjustment consisting of a certain percentile of the relative spread\(^5\) distribution is used for the market “liquidity” risk. The resulting “liquidity” adjusted VaR represents the (say) \(^{99}\text{th}\) percentile base price movement in the asset and the (say) \(^{99}\text{th}\) percentile movement in the relative spread. Figure 2.1 is a graphical representation of VaR with the spread adjustment approach.

\[ \text{Relative Spread} = \frac{\text{Ask price} - \text{bid price}}{\text{Midpoint}} \]

\[ \text{Figure 2.1: VaR with Spread Adjustment. (Source: Bangia et al., 1999)} \]

The model clearly relaxes the unrealistic negligence of spreads in VaR but has several weak points that cannot be overlooked. Firstly, it neglects potential price impacts from its own trading activities, thus implying that it is possible to trade within the spread at all times or equivalently that the own trading volume does not exceed the quoted depth. Furthermore, the severe structural inconsistency of applying spread adjustment to the original value and not to the value after the forecasted loss suggests incorrect timing.

\(^{5}\) The relative spread is defined as [Ask price – bid price]/Midpoint.
Due to this, the model was reformulated to apply the spread adjustment to the forecasted position value at the end of the time horizon. Both models are formally written as follows, initially:

\[ L_{\text{Adj-VaR}} = \text{Mid}_t \{1 - \exp [E(r) - a\theta \sigma_t] \} + \frac{1}{2} \text{Mid}_t (\bar{S} + \bar{a}\bar{\sigma}) . \]  

(2.5.0)

Then,

\[ L_{\text{Adj-VaR}} = \text{Mid}_t \{1 - \exp [E(r) - a\theta \sigma_t] \} + \frac{1}{2} \text{Mid}_t \exp [E(r) - a\theta \sigma_t] . (\bar{S} + \bar{a}\bar{\sigma}) . \]  

(2.5.1)

where, \( \text{Mid}_t \) is the midpoint quote at time \( t \). \( E(r) \) is expected log return. \( \bar{S} \) is average relative speed. \( \bar{a} \) is the \( q \) percentile of the relative spread distribution. \( \bar{\sigma} \) is standard deviation of the relative spread distribution. \( \alpha \) is the \( q \) percentile of the midpoint quote log return distribution. \( \theta \) is correction factor for eventual fat tails of return distribution. \( \sigma_t \) standard deviation of the midpoint quote log return distribution (Subscript \( t \) indicates that in order to capture volatility clustering and time varying volatility, the authors suggest the use of exponentially weighted moving average (EWMA) of past returns. GARCH could be employed as well but can lead to estimation problems for large portfolios).

**2.6 Summary**

This chapter has been dedicated to the review of various models and approaches formulated so far in measuring liquidity risk mostly in the high-frequency finance domain. There are still quite a lot of developments in this area of study that are not captured in this chapter. One other relevant literature worth mentioning is by A. J. Kim [1], where his formulation was derived using a parametric approach by modifying the Weibull distribution and solving the resultant
equation yielding \( \log(\tau) \) as measure for market liquidity. Common to all the models discussed are the elements of spread and volume (market depth). Same is evident in the model I have adopted. The next chapter will be focused on the model adopted to estimate liquidity in my dissertation.
3. METHODOLOGY

The main goal of this paper is to estimate liquidity risk from high frequency tick data with the adjunct task of reconstructing the order book from raw data collected from the London Stock Exchange’s (LSE) Stock-exchange Electronic Trading Service (SETS) platform. While the task of reconstructing the order book cannot be trivialized, it shall be briefly explained but with more emphasis placed on the approach used to estimate liquidity risk from the derived order book.

3.1 Order Book Reconstruction

The primary goal of reconstructing the order book is to regenerate the order book by inferring from captured transaction records, the state of the order book as it appeared *ex ante*. To be able to achieve this, the open source SETS Order Book Rebuilder application by M K Nguyen, N Rayner and S Phelps [] was employed with minimal adjustments. The application was developed using C# with a Microsoft SQL Server database.

The LSE SETS provides three sets of data in Comma Separate Values (CSV) format viz:

*OrderDetails, OrderHistory, TradeReports.*

*OrderDetails* contains information of new orders entering the order book. The most important attributes are: price, volume, time and date, order code, buy or sell indicator.

*OrderHistory* records the history of changes of each order. There are five events that determine the history of an order: *the expiry of an order, the deletion of an order, amendments of its quantities, a partial matching of an order, and a full matching of an order*. Once an order is matched
(fully or partially) the order code of the counter order is also recorded in `OrderHistory` and the details of the transaction are recorded in `TradeReports` (all trades occurring during the auction process are recorded in `TradeReports`). Note that any *non-persistent* order, that is orders such as market orders that are never queued on the order book, are not explicitly recorded in `OrderDetails` or `OrderHistory`. The following steps in Table 3.1 outline the order book reconstruction process:

<table>
<thead>
<tr>
<th>S/N</th>
<th>Task</th>
<th>Processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Data loading</td>
<td>Load LSE SETS data into Microsoft SQL server.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- <code>OrderDetails.csv</code></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- <code>OrderHistory.csv</code></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- <code>TradeReports.csv</code></td>
</tr>
<tr>
<td>2.</td>
<td>A. Retrieve information on all non-persistent orders and missing events from the original LSE data</td>
<td>Add missing Market Orders to <code>OrderDetail</code>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Add missing Market Orders to <code>OrderHistory</code>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Add missing Limit Orders to <code>OrderHistory</code>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Correct aggregate size matches and partial matches.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Identify IceBerg orders in <code>OrderDetail</code> and <code>OrderHistory</code>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Match partial matches as full matches.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Correct the aggregate in Delete &amp; Expire rows.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Populate <code>OrderBook</code> table.</td>
</tr>
<tr>
<td>3.</td>
<td>A. Retrieve complete historic event list, with all events sequenced by their actual occurrence</td>
<td>Simulate and generate Order Book</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Order book outputs in CSV file which will be used for analysis in MATLAB.</td>
</tr>
</tbody>
</table>

Table 3.1: Outline of the SETS Simulator Order Book reconstruction process.
3.2 The New Liquidity Formalism

Acerbi and Scandolo [5][6][7] proposed a new formalism for assessing liquidity risk. The approach taken by this new formalism differ from the models discussed in chapter 2 of this paper, such that, it does not seek to extend any existing liquidity assumptions or hypotheses nor centered on the bid-ask spread or market depth as many others have, but to set a generalized framework for measuring liquidity risk. While the major emphasis of the formalism is centered on measuring portfolio liquidity risk, it does not provide a sufficient basis to cater for high-frequency trading. However, it does set precedence by presenting a standard for valuing any asset using the market data function also referred to as the Marginal Supply Demand Curve (MSDC).

3.3 Marginal Supply Demand Curve (MSDC)

The MSDC is a market function of bid (buy) and ask (sell) prices of an asset, showing the trade-off between the supply and demand for that asset. Our asset in this case is a stock traded on the floor of the LSE SETS and can be traded in standardized units (sizes, quantities or volume). We make some assumptions which are true for assets traded in an order driven market:

- Our asset is quoted by a series of bids and asks prices.
- That each bid and ask is associated with a maximum trading volume.
- That bid prices are always lower than ask prices.

Acerbi and Scandolo [7] represent these market prices mathematically using the MSDC function. To explain the MSDC, consider all the available quotes $m_i$ for an asset $m$, where $s_i$ denotes the trading sizes for $m$. Let us consider a sale of $s_i$ unit of $m$ as $s > 0$, and a purchase of $|s_i|$ unit of $m$. 
as \( s < 0 \). We make not provision for 0 or mid-price as these are not obtainable in any market.

Suppose we want to sell a certain volume \( x \) of asset \( m \), where \( x \leq s \). We expect the best or highest price from the list of bid prices for \( m \) to be to satisfy our sale order, such that these prices are sorted in the order \( m_i > m_{next-i} \) i.e. a decreasing list. Our selling price should be satisfied by \( m_i \) which becomes our effective \( \max(m) \). Alternatively, if we want to buy \( x \) volume of \( m \), where \( x \leq s \), we expect our best ask or offer price to be the \( \min(m) \) sorted in increasing order.

The proceeds from our purchase or sale will become \( \sum_i m_i x_i \). By this, we can define our Asset \( m \), as a good traded in a market characterized by a Marginal Supply Demand Curve (MSDC) which is defined as a map \( m : \mathbb{R}\{0\} \to \mathbb{R} \) satisfying the following conditions:

1. \( m(s_1) \geq m(s_2) \) if \( s_1 < s_2 \). This implies that \( m(s) \) is a decreasing function.

2. \( m \) is càdlàg for \( s < 0 \) (i.e. right-continuous with left limits) and làdcàg for \( s > 0 \) (i.e. left-continuous with right limits).

We refer to \( m^+ := m(0^+) \) the ‘best bid’ and \( m^- := m(0^-) \) the ‘best ask’, and \( \delta m \) as the ‘bid-ask spread’, being the difference between the best ask and the best bid i.e. \( \delta m := m^- - m^+ \geq 0 \).

An asset can assume a constant, positive or negative MSDC function. The nature of the asset determines what type of MSDC function it is. For example, cash has a constant MSDC function due to its perfectly liquid nature. Cash can be defined as an asset representing the currency paid or received when trading another asset. Other assets such as stocks, bonds, or commodities have a positive MSDC function. They are classified as security assets. Assets such as swaps (e.g. interest rate swap, CDS, repo transactions) take on both positive and/or negative MSDC functions. Having established that assets can be represented by an MSDC function. I consider a
simple stock portfolio to be able to adopt the axioms of Acerbi and Scandolo’s [7] formalism on a portfolio of asset(s).

Figure 3.1 shows a typical order book for a stock asset with corresponding MSDC for the stock.

<table>
<thead>
<tr>
<th>Shares</th>
<th>Price($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asks</td>
<td></td>
</tr>
<tr>
<td>2206</td>
<td>1300</td>
</tr>
<tr>
<td>2203</td>
<td>800</td>
</tr>
<tr>
<td>2203</td>
<td>785</td>
</tr>
<tr>
<td>2203</td>
<td>400</td>
</tr>
<tr>
<td>2203</td>
<td>400</td>
</tr>
<tr>
<td>Bids</td>
<td></td>
</tr>
<tr>
<td>2192</td>
<td>400</td>
</tr>
<tr>
<td>2192</td>
<td>1000</td>
</tr>
<tr>
<td>2191</td>
<td>2000</td>
</tr>
<tr>
<td>2189</td>
<td>300</td>
</tr>
<tr>
<td>2188</td>
<td>8550</td>
</tr>
</tbody>
</table>

Order Book of top 5 AZN orders

<table>
<thead>
<tr>
<th>Shares</th>
<th>Price($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asks</td>
<td></td>
</tr>
<tr>
<td>[-11018, -8812)</td>
<td>1300</td>
</tr>
<tr>
<td>[-8812, -6609)</td>
<td>800</td>
</tr>
<tr>
<td>[-6609, -4406)</td>
<td>785</td>
</tr>
<tr>
<td>[-4406, -2203)</td>
<td>400</td>
</tr>
<tr>
<td>[-2203, 0)</td>
<td>400</td>
</tr>
<tr>
<td>Bids</td>
<td></td>
</tr>
<tr>
<td>(0, 2192]</td>
<td>400</td>
</tr>
<tr>
<td>2192, 4384]</td>
<td>1000</td>
</tr>
<tr>
<td>4384, 6575]</td>
<td>2000</td>
</tr>
<tr>
<td>6575, 8764]</td>
<td>300</td>
</tr>
<tr>
<td>8764, 10952]</td>
<td>8550</td>
</tr>
</tbody>
</table>

Table 3.2: List of prices and MSDC for a stock.

### 3.4 The New Portfolio Framework

A key theme central to the new portfolio framework by Acerbi and Scandolo [7] is the rethink of a portfolio as opposed to the general idea set by the Modern Portfolio Theory (MPT) propounded by Markowitz. The MPT defines a portfolio as an algebraic summation of assets give by:

\[ \mathbf{p} = \sum_{i}^{n} \omega_{i}A_{i} \]  

(3.4.1)

The argument by Acerbi and Scandolo maintains that equation (3.3.1) betrays the hidden assumption of perfect market liquidity and will be nonsensical in an illiquid market.
Furthermore, the basic concept of a portfolio connotes a collection of assets i.e. a set notation given by:

$$\mathbf{p} = \{\{A_1 ... A_1\omega_1, \{A_2 ... A_2\omega_2, \ldots, \{A_n ... A_n\omega_n\right\}\}$$

Hence, the value of a portfolio is different from the portfolio itself – the major idea which the MPT obliterates. It is noteworthy to mention at this juncture that, in an illiquid market it is fundamental to realize that both the value and the risk of a given portfolio may take on different values in the hands of two distinct investors. See Appendix [1] for explanation. Having given this background, we define a Portfolio as a vector of real numbers, $\mathbf{p} := (p_0, \mathbf{p}) = (p_0, p_1, \ldots, p_N) \in \mathbb{R}^{N+1}$ where $p_i$ ($i = 0, 1, \ldots, N$) are the holding volume of asset $A_i$, $p_0$ is the holding volume of cash, which is called the portfolio liquidity constant for any portfolio, and $\mathbf{p} = (p_1, \ldots, p_N)$ is the asset’s position referred to as “long-“, “short-“ or “zero-“ positions in asset $A_k$ if $p_k > 0$, $p_k < 0$ or $p_k = 0$ respectively. Also, we denote $\mathcal{P} := \mathbb{R}^{N+1}$ as the portfolio space where the usual operations of addition and scalar multiplication are valid only in $\mathcal{P}$ but not $A$. So, we can write a portfolio $\mathbf{p}$ plus $a$ units of cash asset as $\mathbf{p} + a = (p_0 + a, \mathbf{p})$ for simplicity.

The next step is the valuation of a portfolio from which we can infer liquidity and also estimate the cost of liquidity from. We introduce the term Mark-to-Market (MtM). Marking to market means – “if you want to assess the value of your portfolio you have to confront it with true prices, which are the only physical reality of the market”. We consider two scenarios of marking to market.

1. The Liquidation Mark-to-Market, $L(\mathbf{p})$ value of a portfolio $\mathbf{p} \in \mathcal{P}$ is the sum of each proceeds $P_i$ for asset $A_i$, given by
\[ L(p) := \sum_{i=0}^{N} P_i(p_i) = p_0 + \sum_{i=0}^{N} \int_{0}^{p_i} m_i(x)dx \quad (3.4.3) \]

\( L(p) \) is the total cash we get from the liquidation of all our positions. This situation can be seen as an extreme case where we have to immediately close all positions in our portfolio. This value can be considered as an extremely prudent MtM policy for a portfolio. Such prudence is only necessary for a portfolio subject to external agents that could instigate a sudden closure of all positions. The opposite extreme case is to keep our portfolio as it is, i.e., to liquidate nothing but just marking all illiquid (i.e., non-cash) assets to the best bid price or to the best ask price, depending on whether a long or short position was taken. This leads to the second MtM scenario.

2. The Uppermost Mark-to-Market, \( U(p) \) value of a portfolio \( p \in \mathcal{P} \) is given by

\[ U(p) := \sum_{i=0}^{N} (m_i^+ \cdot \max(p_i, 0) + m_i^- \cdot \min(p_i, 0)) = p_0 + \sum_{i=0}^{N} (m_i^+ \cdot \max(p_i, 0) + m_i^- \cdot \min(p_i, 0)) \quad (3.4.4) \]

where \( m_i^+ \) and \( m_i^- \) are the best bid and best ask prices for asset \( A_i \). The uppermost MtM value can be viewed as the value of a portfolio for an investor who has no cash demands. In this sense, the portfolio has no constraints. This liquidation scenario is an ideal one. The prudential basis for such liquidation is not substantial as it does not probe the market depth at all. From the MSDC we can observe the decreasing function: \( U(p) \geq L(p) \). The difference between \( U(p) \) and \( L(p) \) is termed the Liquidation Cost of a portfolio defined as:

\[ C(p) := U(p) - L(p) \quad (3.4.5) \]

This derivation of \( C(p) \) by Acerbi and Scandolo is essentially fundamental to the estimation of liquidity risks. I adopt the \( C(p) \) function as the primary measure of liquidity in my dissertation.
Further from the liquidation cost function $C(p)$, is the Liquidity Policy $L(p)$ which the formalism introduces as a set(s) of constraints on a portfolio, see Appendix [2] for details.

The resulting measure of risk after the introduction of a $L(p)$ is a convex optimization problem. I did not go beyond the $C(p)$ function to estimate the liquidity risk of an asset or portfolio in my dissertation due to the time constraints required to probe and solve the convex optimization problem resulting from $L(p)$. The formalism also proposes a set of properties for functions $L$, $U$ and $C$ to be acceptable. These properties are specified in Appendix [3]. I later compare my results based on these properties. The next chapter presents the data used for my analysis.
4. DATA

The data employed in this study are obtained from the LSE SETS trading platform. The SETS\textsuperscript{6} is an electronic order driven trading platform. It offers a double auction mechanism for matching buy (bid) or sell (ask) prices in an order book. Figure 4.1 shows the three market operation states of the LSE SETS.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.1.png}
\caption{LSE SETS market operation.}
\end{figure}

The market opens at 7:50 and a random time between 8:00 and 8:00:30 commences an opening auction in which limit orders and market orders are entered and deleted on the order book, however execution is not carried out yet. The purpose of the opening auction is to discover an opening price for the real trading period, which commences randomly after the opening auction.

The trading for the day ends with a closing auction at 16:30 and lasts for five minutes plus a random time up to 30 seconds to ensure a high quality closing price for the day.

\textsuperscript{6} http://www.londonstockexchange.com/products-and-services/trading-services/trading-services.htm
Three sets of data are obtainable from SETS as described in section (3.1). Figure 4.2 shows a schematic representation of these data and the processes involved in finally deriving the order book.

Figure 4.2: Order Book Data generation process.
The order book generated from the SETS Order Book Rebuilder is imported into MATLAB where it is used for liquidity risk analysis.

Two stocks traded in the LSE have been considered for analysis – Astrazeneca PLC (ticker name: AZN) and British American Tobacco PLC (ticker name: BATS). The sample period for both stocks is 2nd June, 2008. There are three order book limit order snapshots taken at 08:30, 12:30 and 16:30 for sampling. The volume is given in number of shares and all orders are time stamped to seconds. Prices are supposedly denominated in US Dollars ($). A sample snapshot of the order book for both stocks is seen in Appendix [4]. From this data, the MSDC for each stock was derived. Also, the liquidation cost for some liquidation scenarios was estimated. The next chapter presents the results obtained.
5. ANALYSIS AND EMPIRICAL RESULTS

In this chapter, the MSDC of the two selected stocks are analyzed. Also, five possible liquidation scenarios are tested with hypothetical quantities calibrated to the order book of each of the stocks. No further assumptions are considered beyond this point, apart from those already stated previously in chapter 3.

5.1 MSDCs of Stocks (AZN and BATS)

The MSDC shows a clear relationship of bids and asks in an order book. At three different snapshots of the order book, the MSDC shows the latent market liquidity. Evident in the MSDC snapshots for each stock is the relative effect of the bid-ask spread with regards to tightness. The MSDC at 12:30 for both stocks shows a very tightly spaced bid-ask spread with well clustered prices tightly spaced from each other. At 16:30, the observation is different as the bid-ask spread widens and the clustered prices observed at 12:30 seem to break apart. This pattern is subsequently confirmed by the high cost of liquidation recorded from liquidating a position at 16:30, which is much lesser when compared to 12:30. Appendix [5] shows the MSDCs for the AZN and BATS stock at 08:30, 12:30 and 16:30.

Also, beyond the bid-ask spread, the MSDC also gives an idea of the available market depth\(^7\), although not precisely. While the tightness of bid-ask spread at 08:30 is relatively as good as the one at 12:30, it can be observed that the order book depth at 12:30 is deeper than 08:30. This is

---

observed by the clustered prices (orders) in the MSDC at 12:30, which should equally sum
higher in volume than 8:30. From this fact, it can be substantiated that the MSDC holds relevant
information useful for analyzing market liquidity at any given time.

5.2 Liquidity Risk Estimation

From (3.4.5), the Liquidation Cost $\mathbf{C}(\mathbf{p})$ of an asset or a portfolio is the measure of the risk
involved in liquidating a market position. This cost of liquidation is actually the dollar value it
would cost a trade position to be liquidated in an illiquid market, thereby paying/attracting an
extra price/cost which is adverse to the safe winding up of the trade. It basically provides a
value to quantity how much risk is involved in a trade position with regards to the market
liquidity. For example, using the order book in Table 3.2, if we sell $\mathbf{p} = 5000$ shares we get:

\[
\mathbf{L}(5000) = 2192.400 + 2192.1000 + 2191.2000 + 2189.300 + 2188.1300
= 10,951,900
\]

\[
\mathbf{U}(5000) = 2192.5000 = 10,960,000
\]

\[
\mathbf{C}(5000) = \mathbf{U}(5000) - \mathbf{L}(5000) = 8,100
\]
is the liquidation cost for 5000 shares.

Also, if we consider a simple portfolio of two assets, such that our portfolio $\mathbf{\hat{p}} = (p_0, p_1, p_2) =$
(10000,3000,−1000). 10000 being the cash asset in our portfolio. We obtain the following by
Marking our Long (3000) and Short (2000) positions to the market:

\[
\mathbf{L}(\mathbf{\hat{p}}) = 10000 + (2192.400 + 2192.1000 + 2191.1600) - (2203.400 + 2203.400 + 2203.200)
= 10000 + 6,574,400 - 2,203,000
= 4,381,400
\]

\[
\mathbf{U}(\mathbf{\hat{p}}) = 10000 + (2192.3000) - (2203.1000)
= 4,383,000
\]
\[ C(\mathbf{p}) = U(\mathbf{p}) - L(\mathbf{p}) = 1,600, \] 
which is the liquidation cost for \( \mathbf{p} \).

Now, to estimate the liquidation costs for the AZN and BATS stocks, the following liquidation scenarios are evaluated:

1. Short 10 various quantities of AZN stock at 12:30 and 16:30.
2. Long AZN stock and Short BATS – 10 possible quantities of each at 12:30 and 16:30.
3. Long AZN stock and Short BATS – 10 x 10 possible quantities of each at 12:30 and 16:30.

I use a vector of 10 pair of sample quantities for both AZN and BATS stocks for all scenarios, given as:

\[
\begin{align*}
\text{AZN} & = [1000 \ 2000 \ 3000 \ 4000 \ 5000 \ 10000 \ 20000 \ 30000 \ 40000 \ 50000] \\
\text{BATS} & = [500 \ 1000 \ 1500 \ 2000 \ 2500 \ 5000 \ 10000 \ 15000 \ 20000 \ 25000]
\end{align*}
\]

Also,

Cash/riskless asset in the portfolio (for scenarios 2 to 3) = 50000.

Results show that, liquidation costs for the AZN stock at 16:30 are higher than 12:30 due to the looseness of the bid-ask spread during this time. Also, liquidation costs increases with corresponding increase in volume of stock to be liquidated. However, at some certain quantity limits, cost of liquidation is unchanged. This is because for fewer quantities of stock that does not require much probe into the market depth, the value of \( C(\mathbf{p}) \) is very low or almost zero due to negligible difference between \( U(\mathbf{p}) \) and \( L(\mathbf{p}) \). Appendix [6] shows the graphical and numerical results for the four scenarios.
5.3 Summary

The results obtained from my analysis are consistent with the concave properties of $U(p)$ and $L(p)$, and the convex property of $C(p)$. A further relevance to the liquidation cost obtained from the four scenarios is stylized by the MSDCs derived for each stock. Two facts stylized by the MSDC which are exemplified in the corresponding results obtained from $C(p)$ are:

1. Liquidation costs increases with the volume of stock to be liquidated.
2. Loose or wide bid-ask spread characterizes illiquid markets with impact on liquidation cost.

With these results, the task of estimating liquidity risk from high frequency tick data using the MSDC as a primary tool is very much feasible.
6. CONCLUSION

One of the reasons why liquidity is an elusive market idea is due to the many dimensions in which it is manifested. From the availability of loans at affordable rates to the ability to fund transactions, easy cash mobility, inter-bank loans and more commonly the exchange of an asset in return for cash, liquidity is very evident and essential for markets to thrive. While we can observe liquidity, there is no clear cut approach or method to quantify it. Much worse is the availability of data to begin with. Acknowledging these issues associated with liquidity motivated the use of high frequency market data for this analysis, with the benefit of ready availability of such market data. Beyond data availability is the approach employed in precisely quantifying liquidity itself. Several approaches have been devised by many authors but not too many, if not any, have been able to devise a holistic approach to quantify liquidity and liquidity risk without bias to the market structure in regards.

Two ideas – Marginal Supply Demand Curve (MSDC) and Mark-to-Market (MtM) are the key instruments adopted in my analysis. The MSDC shows a precise picture of liquidity by relating market prices of bids and asks and their corresponding volume to one another, and the trade-off between both sides characterized by a decreasing function. This information is further utilized by the MtM functions posited in Acerbi & Scandolo’s new formalism for liquidity risk. The combination of these two ideas is a starting point towards designing a trading strategy that incorporates liquidity risk. Future work can be tailored in this regard.
7. REFERENCES


[23] M. K. Nguyen, N. Rayner and S. Phelps. Inferring the state of a double-auction market from empirical high-frequency transaction data, 2010


8. **APPENDIX**

1. **Portfolio value example**
   In an illiquid market the value of a portfolio could be valued differently by investors. For instance, let \( p \) be a portfolio of very illiquid bonds with tenor 7-10 years with face value of $1 Million, quoting around par.

   a. Let Alan be an investor who (for some reason) may afford to keep \( p \) until maturity.

   b. Let Ben be an investor who (for some reason) will be obliged to liquidate periodically large portions of \( p \).

   It is clear that,

   1. For Alan the portfolio is worth more or less $1 million. For Ben certainly much less because liquidating he will face a liquidation cost. Ben may very well be willing to sell it immediately for much less than $1 million.

   2. Clearly, for Alan the liquidity risk of \( p \) is zero, whereas for Ben is very large.

2. **Liquidity Policy \( L(p) \)**

   The definitions of the Liquidation MtM value \( L(p) \) and the Uppermost MtM value \( U(p) \) suggests that the value of a portfolio \( p \) is subject to constraints, which represents certain cash constraints an investor should be able to meet by wholly or partially liquidating set trade positions. Constraints to liquidation could vary from cash to trade exposure limits, etc. For example, an investor can impose a cash liquidity policy signifying a readiness to
liquidate set positions that must satisfy the minimum cash requirement stipulated or impose market risk VaR limits on set positions, or credit limits, or capital constraints. All the constraints that an investor imposes can be represented as a subset of the underlying portfolio space \( \mathcal{P} \). These constraints are collectively referred to as a liquidity policy. The liquidity policy has a property of a \textit{closed and convex} subset of the underlying portfolio space which leads to a convex optimization problem not explored in my dissertation. Refer to Acerbi & Scandolo [7] for the exact mathematical definition of a liquidity policy and further explanation.

3. **Properties of functions \( L, U \) and \( C \)**

By definitions of \( L, U \) and \( C \) in (3.4), (Acerbi & Scandolo[7]) give the following properties (Proposition 4.1 in their formalism).

Let \( p, q \in \mathcal{P} \) and \( \theta \in [0, 1] \).

- **The liquidation MtM value operator \( L : \mathcal{P} \to \mathbb{R} \)**
  1. \textit{is concave}, i.e., \( L(\theta p + (1 - \theta)q) \geq \theta L(p) + (1 - \theta)L(q) \).
  2. \( L(\lambda p) \leq \lambda L(p) \) if \( \lambda \leq 1 \).

- **The uppermost MtM value operator \( U : \mathcal{P} \to \mathbb{R} \)**
  1. \textit{is concave}, i.e., \( U(\theta p + (1 - \theta)q) \geq \theta U(p) + (1 - \theta)U(q) \).
  2. \textit{is positive homogeneous}, i.e., \( U(\lambda p) = \lambda U(p) \) if \( \lambda \geq 0 \).

- **The uppermost liquidation cost operator \( C : \mathcal{P} \to \mathbb{R}^+ \)**
  1. \textit{is convex}, i.e., \( C(\theta p + (1 - \theta)q) \leq \theta C(p) + (1 - \theta)C(q) \).
  2. \( C(\lambda p) \geq \lambda C(p) \) if \( \lambda \geq 1 \).
4. Order book snapshots for AZN and BATS stocks at 02/06/2008 08:30

AZN Order Book snapshot at 02/06/2008 08:30

<table>
<thead>
<tr>
<th>Price</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2203</td>
<td>400</td>
</tr>
<tr>
<td>2203</td>
<td>400</td>
</tr>
<tr>
<td>2203</td>
<td>785</td>
</tr>
<tr>
<td>2203</td>
<td>800</td>
</tr>
<tr>
<td>2206</td>
<td>1300</td>
</tr>
<tr>
<td>2206</td>
<td>414</td>
</tr>
<tr>
<td>2207</td>
<td>2000</td>
</tr>
<tr>
<td>2208</td>
<td>500</td>
</tr>
<tr>
<td>2211</td>
<td>540</td>
</tr>
<tr>
<td>2211</td>
<td>1100</td>
</tr>
<tr>
<td>2212</td>
<td>8550</td>
</tr>
<tr>
<td>2212</td>
<td>600</td>
</tr>
<tr>
<td>2218</td>
<td>1270</td>
</tr>
<tr>
<td>2238</td>
<td>10000</td>
</tr>
<tr>
<td>2243</td>
<td>9100</td>
</tr>
<tr>
<td>2244</td>
<td>983</td>
</tr>
<tr>
<td>2250</td>
<td>975</td>
</tr>
<tr>
<td>2256</td>
<td>930</td>
</tr>
<tr>
<td>2380</td>
<td>300</td>
</tr>
<tr>
<td>2420</td>
<td>1000</td>
</tr>
<tr>
<td>2641</td>
<td>125</td>
</tr>
<tr>
<td>3000</td>
<td>700</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Price</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2192</td>
<td>400</td>
</tr>
<tr>
<td>2192</td>
<td>1000</td>
</tr>
<tr>
<td>2191</td>
<td>2000</td>
</tr>
<tr>
<td>2189</td>
<td>300</td>
</tr>
<tr>
<td>2188</td>
<td>8550</td>
</tr>
<tr>
<td>2187</td>
<td>1256</td>
</tr>
<tr>
<td>2184</td>
<td>1000</td>
</tr>
<tr>
<td>2183</td>
<td>1221</td>
</tr>
<tr>
<td>2179</td>
<td>9100</td>
</tr>
<tr>
<td>2170</td>
<td>5000</td>
</tr>
<tr>
<td>2170</td>
<td>25000</td>
</tr>
<tr>
<td>2160</td>
<td>1874</td>
</tr>
<tr>
<td>2154</td>
<td>1818</td>
</tr>
<tr>
<td>2148</td>
<td>1877</td>
</tr>
<tr>
<td>2129</td>
<td>10000</td>
</tr>
<tr>
<td>2000</td>
<td>75</td>
</tr>
<tr>
<td>1780</td>
<td>2000</td>
</tr>
<tr>
<td>1700</td>
<td>500</td>
</tr>
</tbody>
</table>
BATS Order Book snapshot at 02/06/2008 08:30

<table>
<thead>
<tr>
<th>ASKs</th>
<th>Price</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1877</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>1878</td>
<td>595</td>
</tr>
<tr>
<td></td>
<td>1878</td>
<td>500</td>
</tr>
<tr>
<td></td>
<td>1878</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>1878</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>1878</td>
<td>1500</td>
</tr>
<tr>
<td></td>
<td>1880</td>
<td>1588</td>
</tr>
<tr>
<td></td>
<td>1880</td>
<td>514</td>
</tr>
<tr>
<td></td>
<td>1881</td>
<td>425</td>
</tr>
<tr>
<td></td>
<td>1881</td>
<td>2712</td>
</tr>
<tr>
<td></td>
<td>1883</td>
<td>1400</td>
</tr>
<tr>
<td></td>
<td>1885</td>
<td>9800</td>
</tr>
<tr>
<td></td>
<td>1887</td>
<td>800</td>
</tr>
<tr>
<td></td>
<td>1893</td>
<td>620</td>
</tr>
<tr>
<td></td>
<td>1895</td>
<td>1225</td>
</tr>
<tr>
<td></td>
<td>1896</td>
<td>10432</td>
</tr>
<tr>
<td></td>
<td>1905</td>
<td>10500</td>
</tr>
<tr>
<td></td>
<td>1907</td>
<td>15897</td>
</tr>
<tr>
<td></td>
<td>1910</td>
<td>4000</td>
</tr>
<tr>
<td></td>
<td>1919</td>
<td>1143</td>
</tr>
<tr>
<td></td>
<td>1920</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>1923</td>
<td>1157</td>
</tr>
<tr>
<td></td>
<td>1928</td>
<td>4000</td>
</tr>
<tr>
<td></td>
<td>1928</td>
<td>1100</td>
</tr>
<tr>
<td></td>
<td>1929</td>
<td>15000</td>
</tr>
<tr>
<td></td>
<td>1950</td>
<td>1060</td>
</tr>
<tr>
<td></td>
<td>1959</td>
<td>11352</td>
</tr>
<tr>
<td></td>
<td>1970</td>
<td>2000</td>
</tr>
<tr>
<td></td>
<td>1990</td>
<td>939</td>
</tr>
<tr>
<td></td>
<td>2050</td>
<td>1000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BIDs</th>
<th>Price</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1875</td>
<td>666</td>
</tr>
<tr>
<td></td>
<td>1875</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>1874</td>
<td>390</td>
</tr>
<tr>
<td></td>
<td>1872</td>
<td>2711</td>
</tr>
<tr>
<td></td>
<td>1872</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>1870</td>
<td>2000</td>
</tr>
<tr>
<td></td>
<td>1869</td>
<td>9800</td>
</tr>
<tr>
<td></td>
<td>1868</td>
<td>1229</td>
</tr>
<tr>
<td></td>
<td>1865</td>
<td>1246</td>
</tr>
<tr>
<td></td>
<td>1856</td>
<td>10432</td>
</tr>
<tr>
<td></td>
<td>1855</td>
<td>10500</td>
</tr>
<tr>
<td></td>
<td>1852</td>
<td>2135</td>
</tr>
<tr>
<td></td>
<td>1852</td>
<td>250</td>
</tr>
<tr>
<td></td>
<td>1850</td>
<td>15897</td>
</tr>
<tr>
<td></td>
<td>1847</td>
<td>2184</td>
</tr>
<tr>
<td></td>
<td>1843</td>
<td>2314</td>
</tr>
<tr>
<td></td>
<td>1835</td>
<td>15000</td>
</tr>
<tr>
<td></td>
<td>1832</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>1805</td>
<td>11352</td>
</tr>
<tr>
<td></td>
<td>1744</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>1650</td>
<td>600</td>
</tr>
</tbody>
</table>
5. **Marginal Supply and Demand Curves for AZN and BATS stocks.**

MSDC for AZN at three different snapshots of the order book (08:30, 12:30, 16:30)

Bid-ask spread tightness can be seen from the 3 various snapshots. Strongest at 12:30, then at 08:30 and weakest at 16:30.
MSDC for BATS at three different snapshots of the order book (08:30, 12:30, 16:30)

Bid-ask spread tightness can be seen from the 3 various snapshots. Strongest at 12:30, then at 08:30 and weakest at 16:30.
6. Liquidation scenarios for AZN and BATS stocks.

\[ \text{AZN} = [1000 \ 2000 \ 3000 \ 4000 \ 5000 \ 10000 \ 20000 \ 30000 \ 40000 \ 50000] \]

\[ \text{BATS} = [500 \ 1000 \ 1500 \ 2000 \ 2500 \ 5000 \ 10000 \ 15000 \ 20000 \ 25000] \]

RISKLESS CASH ASSET = 50,000

A. Short 10 various quantities of AZN stock at 12:30 and 16:30.
Liquidation cost for AZN at 12:30
[0 0 0 0 404 6412 43250 106009 222781 417633]

Liquidation cost for AZN at 16:30
[15444 59444 103444 147444 191444 411444 895883 1609433 2349127 3119127]

Results show higher liquidation cost at 16:30 due to the looseness of bid ask spread as seen from the MSDC.

B. Long AZN stock and Short BATS – 10 possible quantities of each at 12:30 and 16:30.
Liquidation cost for LONG AZN & SHORT BATS Stock at 12:30
[0 0 0 0 404 6663 49845 122604 253757 470233]

Liquidation cost for LONG AZN & SHORT BATS Stock at 16:30
[15444 59444 103444 147444 191444 411444 964483 1748033 2635871 3613752]

Results show higher liquidation cost at 16:30 due to the looseness of bid ask spread as seen from the MSDC. At 12:30, region where the curve overlaps the x-axis show little or minimal liquidity risk involved.

C. Long AZN stock and Short BATS – 10x10 possible quantities of each at 12:30 and 16:30.

![Graph showing liquidation cost for Long AZN and Short BATS stocks]  

The liquidation cost $C(p)$ is a convex shape obeying the convex property. $U(p)$ and $L(p)$ appear concave in pattern which also obeys the concave property. This information reflects the minimum liquidation cost to the maximum obtainable from a 10x10 pair of both AZN and BATS stocks. Also, the shape of $C(p)$ is semi-convex due to the one sided combination of Long: AZN and Short: BAT stocks. A full combination would have 10x10x10 of both stocks i.e by Long: BATS and Short: AZN conversely and that would result to a perfectly convex $C(p)$. 


Liquidation MtM $U(p)$ function for MSDC showing a concave pattern

Liquidation MtM $L(p)$ function for MSDC showing a concave pattern
7. **Program Listing.**

**MATLAB CODES. – 1 (mdscMain.m): The main script.**

```matlab
%% Order Book Analysis.
% Extraction of Marginal Supply Demand Curve (MSDC) data from the order book.
% MSDC data will be 2-Matrices for the Bids & Ask

obStatesAsk = find(batAsk == 0); % Order Book States for Ask side. 199,998 States of OB
obStatesBid = find(batBid == 0); % Order Book States for Bid side. 199,830 States of OB.

% Get order book states.
% A2N Ask States.
% Index to find are derived from scraps.m file for 08:30, 12:30 and 16:30.
aznA1a = find(obStatesAsk == 88) % 8:30 = 4
aznA1b = find(obStatesAsk == 956141) % 12:30 = 25532
aznA1c = find(obStatesAsk == 2178253) % 16:30 = 52408

% A2N Bid States
aznB1a = find(obStatesBid == 95) % 8:30 = 5
aznB1b = find(obStatesBid == 839671) % 12:30 = 25538
aznB1c = find(obStatesBid == 1848522) % 16:30 = 52410

% BATS Ask States
batA1a = find(obStatesAsk == 88) % 8:30 = 3
batA1b = find(obStatesAsk == 887051) % 12:30 = 18887
batA1c = find(obStatesAsk == 1950326) % 16:30 = 38524

% BATS Bid States
batB1a = find(obStatesBid == 23) % 8:30 = 1
batB1b = find(obStatesBid == 760045) % 12:30 = 18932
batB1c = find(obStatesBid == 1686143) % 16:30 = 38544

% Get OB snapshots for 08:30, 12:30, 16:30
% A2N
% 08:30
aznAsk = (89:110,1:2) % OB snapshot @ 92
aznBid = (96:113,1:2) % OB snapshot @ 96

% 12:30
aznAsk = (956142:956181,1:2) % OB snapshot @ 956163
aznBid = (839672:839715,1:2) % OB snapshot @ 839701

% 16:30
aznAsk = (2178254:2178278,1:2) % OB snapshot @ 2178254
aznBid = (1848523:1848541,1:2) % OB snapshot @ 1848525

% BATS
% 08:30
batAsk = (96:126,1:2) % OB snapshot @ 92
batBid = (24:46,1:2) % OB snapshot @ 24

% 12:30
batAsk = (887052:887101,1:2) % OB snapshot @ 887063
batBid = (760046:760088,1:2) % OB snapshot @ 760060

% 16:30
batAsk = (1950327:1950365,1:2) % OB snapshot @ 1950332

```
batBid(1686144:1686172,1:2) % OB snapshot @ 1686144

%% MSDCs for AZN and BAT
%% Visualize Order Book
%% AZN
aznPlot2(89,110,96,113) % @ 08:30
aznPlot2(956142,956181,839672,839715) % @ 12:30
aznPlot2(2178254,2178278,1848523,1848541) % @ 16:30

subplot(2,2,1), aznPlot2(89,110,96,113)
subplot(2,2,2), aznPlot2(956142,956181,839672,839715)
subplot(2,2,3), aznPlot2(2178254,2178278,1848523,1848541)

%% BATS
batPlot2(96,126,24,46) % @ 08:30
batPlot2(887052,887101,760046,760088) % @ 12:30
batPlot2(1950327,1950365,1686144,1686172) % @ 16:30

subplot(2,2,1), batPlot2(96,126,24,46)
subplot(2,2,2), batPlot2(887052,887101,760046,760088)
subplot(2,2,3), batPlot2(1950327,1950365,1686144,1686172)

%% Liquidity risk estimation.

% 1 Quantity. SHORT AZN and SHORT BATS
assetSizeAZN = 20000;
assetSizeBAT = 20000;
obState1 = 25538;        % OB State of AZN Bid side @ 12:30
obState2 = 18932;        % OB State of BAT bid side @ 12:30
obStatea = 52410;        % OB State of AZN Bid side @ 16:30
obStateb = 38544;        % OB State of BAT bid side @ 16:30
riskless = 50000;

[lowVal1 lowVal2] = lowerMtm(assetSizeAZN,assetSizeBAT, obStatea, obStateb)
% L lowVal1 - AZN, lowVal2 - BAT
[upVal1 upVal2] = upperMtm(assetSizeAZN, assetSizeBAT, obStatea, obStateb)
% U upVal1 - AZN, upVal1 - BAT

costOfLiq_1 = upVal1 - lowVal1 % Cost of Liquidity for AZN 43250, 895883
costOfLiq_2 = upVal2 - lowVal2 % Cost of Liquidity for BAT 71225, 21296
% From above, the cost of liquidation is higher @ 16:30 than 12:30 due
% reduction of liquidity.

% 10 different Quantities. SHORT AZN and SHORT BATS - SINGLE
% if assetSizes should be varied ... then lets see the effect of C

assetSizeAZN2 = [1000 2000 3000 4000 5000 10000 20000 30000 40000 50000];
assetSizeBAT2 = [500 1000 1500 2000 2500 5000 10000 15000 20000 25000];
for i = 1:numel(assetSizeAZN2)
    [lowValue1(i) lowValue2(i)] =
        lowerMtm(assetSizeAZN2(i),assetSizeBAT2(i), obStatea, obStateb);
    [upValue1(i) upValue2(i)] = upperMtm(assetSizeAZN2(i), assetSizeBAT2(i), obStatea, obStateb);
    aznliquidityCost(i) = upValue1(i) - lowValue1(i); %
    batliquidityCost(i) = upValue2(i) - lowValue2(i); %
Estimating Liquidity Risk from High Frequency Tick Data

```matlab
% Portfolio Liquidity Cost
portliquidityCost(i) = (upValue1(i) + upValue2(i)) - (lowValue1(i) + lowValue2(i));

% Liquidation cost for just one stock of AZN or BATS
plot(assetSizeAZN2, aznliquidityCost); plot(assetSizeBAT2, batliquidityCost);
end

%% LONG ONE STOCK AND SHORT ONE STOCK CASE

% LONG AZN SHORT BATS simple case.
% 1 portfolio
obState1a = 25532;        % OB State of AZN Ask side @ 12:30
obState1b = 25538;        % OB State of AZN Bid side @ 12:30
obStatea1 = 52408;        % OB State of AZN Ask side @ 16:30
obStatea2 = 52410;        % OB State of AZN Bid side @ 16:30
obState2a = 18887;        % OB State of BAT Ask side @ 12:30
obState2b = 18932;        % OB State of BAT Bid side @ 12:30
obStateb1 = 38524;        % OB State of BAT Ask side @ 16:30
obStateb2 = 38544;        % OB State of BAT Bid side @ 16:30
upper = upperMtmBS(20000, 1000, obStatea1, obStatea2, obStateb1, obStateb2, 50000, 0) % U with LONG AZN and SHORT BAT
lower = lowerMtmBS(20000, 1000, obStatea1, obStatea2, obStateb1, obStateb2, 50000, 0) % L with LONG AZN and SHORT BAT
c1 = upper - lower % 43,250 @ 12:30; 895,883 @ 16:30

% 10 Possible Liquidation scenarios of LONG AZN SHORT BATS.
% Loop through for various pair sizes of AZN and BAT with 50,000 cash
% as riskless asset SINGLE

assetSizeAZN2 = [1000 2000 3000 4000 5000 10000 20000 30000 40000 50000];
assetSizeBAT2 = [500 1000 1500 2000 2500 5000 10000 15000 20000 25000];
for j = 1:numel(assetSizeAZN2)
    lowVal(j) = lowerMtmBS(assetSizeAZN2(j), assetSizeBAT2(j), obStatea1, obStatea2, obStateb1, obStateb2, 50000, 0);
    upVal(j) = upperMtmBS(assetSizeAZN2(j), assetSizeBAT2(j), obStatea1, obStatea2, obStateb1, obStateb2, 50000, 0);
    portliquidityCost(j) = upVal(j) - lowVal(j); % Portfolio Liquidity Cost:
end
plot(portliquidityCost)
```

% 10x10 Possible Liquidation scenarios of LONG AZN SHORT BATS.
% Loop through for various pair sizes of AZN and BAT with 50,000 cash
% as riskless asset MULTIPLE
for j = 1:numel(assetSizeAZN2)
    for k = 1:numel(assetSizeBAT2)
        low2Val(j,k) = lowerMtmBS(assetSizeAZN2(j), assetSizeBAT2(k),
                                 obStatea1, obStatea2, obStateb1, obStateb2, 50000, 0);
        up2Val(j,k) = upperMtmBS(assetSizeAZN2(j), assetSizeBAT2(k),
                                 obStatea1, obStatea2, obStateb1, obStateb2, 50000, 0);
        c(j,k) = up2Val(j,k) - low2Val(j,k); %Portfolio Liquidity Cost : C
    end
end

% Plots
surf1(assetSizeAZN2,assetSizeBAT2,c)
surf1(assetSizeAZN2,assetSizeBAT2,up2Val)
surf1(assetSizeAZN2,assetSizeBAT2,low2Val)

Other scripts, source codes and data are included in the CD attached.