

# The Dynamics of Ex-ante High-Frequency Liquidity: An Empirical Analysis<sup>\*</sup>

Georges Dionne<sup>a,b,†</sup>      Xiaozhou Zhou<sup>b,c</sup>

<sup>a</sup>Canada Research Chair in Risk Management, HEC Montréal, Montréal H3T 2A7, Canada

<sup>b</sup>Finance Department, HEC Montréal, Montréal H3T 2A7, Canada

<sup>c</sup>Faculty of Management (ESG), University of Quebec at Montreal (UQAM), H3C 3P8, Canada

## Abstract

Using tick-by-tick data and the reconstructed open LOB data from the Xetra trading system, we investigate the impact of trade duration, quote duration and other exogenous variables on ex-ante liquidity embedded in an open LOB. Our modeling involves decomposing the joint distribution of the ex-ante liquidity measure into simple and interpretable distributions. The decomposed factors are Activity, Direction and Size. Our results suggest that trade durations and quote durations do influence the ex-ante liquidity changes. Short-run variables, such as spread change and volume, also predict the probability of liquidity changes.

Keywords: Decomposition model, Limit order book, Xetra Liquidity Measure (XLM), Ex-ante liquidity, LogACD process.

JEL classification: C22 C41 C53 G11

---

<sup>\*</sup>We thank Yann Bilodeau for his help in constructing the dataset and comments. We also thank Christian Gouriéroux, Diego Amaya, Gabriel Yergeau, Maria Pacurar, Tolga Cenesizoglu for their remarks. Georges Dionne acknowledges financial support from the Social Sciences and Humanities Research Council (SSHRC) in Canada and the Canada Foundation for Innovation, and Xiaozhou Zhou thanks Fonds de Recherche sur la société et la culture de Québec (FRQSC), the Canada Research Chair in Risk Management, and Centre interuniversitaire sur le Risque, les Politiques, Economiques et l'Emploi (CIRPEE) for financial support.

<sup>†</sup>Corresponding author at: Canada Research Chair in Risk Management, HEC Montréal, 3000 Côte-Sainte-Catherine, Montreal, Canada. H3T 2A7. Tel.: +1 514 340 6596. E-mail addresses: georges.dionne@hec.ca (G. Dionne), xiao-zhou.zhou@hec.ca (X. Zhou).

# 1 Introduction

Liquidity is a complex concept and has been one of the most important issues in financial research for a long time. The term 'liquidity' is interpreted differently by various market participants. Market regulators see liquidity as either the capacity to buy and sell a large quantity of financial securities, or as the total turnover of assets in a given time interval. For individual traders, the level of liquidity relates more to the quantity available when changing their positions at buying or selling side. Nowadays, liquidity evolves with the development of high-frequency trading (HFT) <sup>1</sup>. As shown by Hendershott and Riordan (2013), for stocks of DAX30, high-frequency traders present 52% of market order volume and 64% of nonmarketable limit order volume. Understanding how liquidity behaves in this HFT environment is essential for all buy-side traders<sup>2</sup> and market regulators. However, there are few quantitative analysis of the ex-ante liquidity in an ultra high frequency environment. The objective of this paper is twofold. First we model the dynamic of ex-ante liquidity, which is offered by the Limit Order Book (LOB), from a general view through the analysis of available electronic data in the Xetra open LOB system. Second, we investigate the effect of exogenous high-frequency variables on this ex-ante liquidity measure.

Ex-ante liquidity is embedded in LOB and there already exists a large number of theoretical microstructure models emphasizing the state of LOB because of its importance in providing market liquidity and discovering price formation. However, there is a huge gap between theoretical and empirical research for the analysis of high-frequency trading mechanism and liquidity. This is due to the fact that market order traders and LOB traders behave in a complicated way, and theoretical models do not capture all of these complexities. For example, the LOB traders in Parlour (1998) cannot choose multiple limit-order price strategies; in Foucault (1999), the lifetime of a limit order can not last more than one period; and in Foucault et al. (2005), the limit orders cannot be canceled once they are placed in the open LOB. All these restrictions may have an impact on the ex-ante liquidity provision.

In the market microstructure framework, other than spontaneous supply and demand in the market, the liquidity provision inherently depends on various exogenous factors such as the trading mechanism, the information disclosure process and regulatory issues. For instance, Viswanathan and Wang (2002) show that the slope of the equilibrium bid price is flatter than that in a dealership market since the discriminatory pricing rule intensifies the liquidity provision. The empirical results of Boehmer et al. (2005) suggest that an increase in pre-trade transparency can improve market liquidity. Riordan and Storckenmaier (2012) examine the effect of a technological upgrade on the market liquidity of 98 actively traded German stocks and show that both effective spreads and average price impacts drop with the upgrade. Bloomfield et al. (2015) use

---

<sup>1</sup>In high-frequency trading, traders make profit by using algorithms to quickly process information and detect the trade opportunities.

<sup>2</sup>Buy-side traders includes individuals, mutual funds, financial institutions, etc.

a laboratory market to investigate how the ability to hide orders affects market liquidity and traders' strategies and suggest that liquidity and informational efficiency are not affected when traders adapt their behavior to the different opacity regimes. Another important concern is the availability of data. In previous studies, widely used high-frequency liquidity measures, such as trade impact and effective spread, are obtained from trade-related databases. One problem with these liquidity measures is that they are measured only when there is a trade. With the introduction of an open LOB trading mechanism, the access to more complete high-frequency data becomes possible. As a result, we can use the information in the LOB to compute two different measures of liquidity: ex-ante liquidity and ex-post liquidity. Up to now there exists little consensus on the most appropriate liquidity measure in terms of efficiency and accuracy due to its multi-faceted complexity in both cross-sectional (depth, width, resiliency, etc.) and frequency (high-frequency or low-frequency based) dimensions.

This empirical paper concentrates on the dynamic of the ex-ante liquidity embedded in an open LOB (depth and price impact) and the effect of high-frequency variables on its dynamics <sup>3</sup>. Recent studies related to ex-ante liquidity include Irvine et al. (2000), Coppejans et al. (2004), Domowitz et al. (2005), Giot and Grammig (2006), Beltran-Lopez et al. (2009), and Beltran-Lopez et al. (2011). According to Aitken and Comerton-Forde (2003), ex-post liquidity measures involve trade-based measures, while ex-ante liquidity measures are order-based. The former measures are the most widely used and indicate what the traders have obtained in the realized transaction. The second group captures the cost related to potential immediate trading. This cost is quasi-fixed for traders with a small quantity to trade and is equal to half of the bid-ask spread. In this case we do not need any information about the LOB. However, for the larger traders, the spread will surely underestimate the associated cost when the quantity to trade is larger than the quantity available at first level. This requires a liquidity measure that can allow for volume-related price impact. To this end, we construct and model the ex-ante liquidity measure with different trading volumes (from 5 000 to 2 5000 euros).

Another characteristic of the ex-ante liquidity measure is that it can be computed even when there is no trade. As noted by Beltran-Lopez et al. (2011), in 2004, for the average stock in the DAX30 of Xetra trading system, the average daily number of limit orders is 12,785, i.e. 25 limit orders per minute, whereas the average number of trades is four per minute. The corresponding numbers are 334 per minute for limit orders and 10 per minute for trades in July 2010. As the trend of high frequency trading continues to increase rapidly, market regulators and institutional investors need more continuous measures of liquidity to update their information set. The traditional ex-post liquidity measurement cannot meet this requirement. The temporal relation between updates of LOB and transactions is shown in Figure 1 where transactions and updates of LOB are denoted by circles and squares respectively. As we see the quote updates

---

<sup>3</sup>A large number of empirical studies examine the effect of HFT on the ex-post liquidity (Hendershott et al. (2011), Hasbrouck and Saar (2013), Menkveld (2013), Hendershott and Riordan (2013), Brogaard et al. (2014), among others). Survey articles include Biais and Wooley (2011), Johes (2012), O'Hara (2015).

are usually much more frequent than the trade updates and may then contain more information on liquidity.

[Insert Figure 1 here]

This study models the evolution of ex-ante liquidity measurement using a decomposition model which allows for various factors in a flexible way and investigates the effect of exogenous high frequency variables on this ex-ante liquidity measurement. To our knowledge, we are the first to consider modeling ex-ante liquidity using a decomposition model and including a large set of factors. Given the particularity of UHF data and the complexity of microstructure analysis, one possible modeling framework involves consistently decomposing the joint distribution of a target variable into simple and interpretable distributions. The idea of decomposition was pioneered by Rogers and Zane (1998) and aims at constructing observation-driven models in the sense of Cox et al.(1981). The decomposition model was first used in analyzing transaction price dynamics. Hausman et al. (1992) and Russell and Engle (2005) propose an Autoregressive Conditional Multinomial (ACM) and ordered Probit model respectively. Rydberg and Shephard (2003) manage to achieve the same goals by decomposing the joint distribution of tick-by-tick transaction price changes into three sequential components. The first component is designated 'active' and used to indicate whether the price change will occur. The second component called 'direction', indicates the direction of price change, and the third, 'size', measures the absolute size of the price change. The decomposition model is used to predict price movements or price level with the help of simulations. McCulloch and Tsay (2001) aim at modeling the transaction price change process using the decomposition model. In their framework, they initialize a price change and duration (PCD) model that decomposes the price changes into four factors and introduce time and liquidity dimension in modeling price change dynamics. The duration between two consecutive transactions and the number of trades during this duration are modeled as implicit factors for the price changes. In total, they use six conditional models to capture the dynamic of price changes. Manganelli (2005) applies the decomposition methodology in investigating the simultaneous interaction between duration, volume and return. Two subgroups, classified by the trade intensity, perform different dynamics. The decomposition framework remains flexible for more complicated modeling and, depending on different modeling assumptions, addition or deletion of certain factors is possible.

Our paper differs from the existing literature in several dimensions. First of all, instead of aggregating the time for a fixed interval, our analysis contains a time dimension. Specifically, our paper investigates the role of trade and quote durations in explaining the dynamic of ex-ante liquidity provision. It is widely known that trades may contain private information that will be further incorporated in the quote updates. Hasbrouck (1991) shows that, in an order-driven market, the price impact of trades is positive, and large trades cause the spread to widen. Intuitively, when facing informed traders, market makers protect themselves by quickly widening the spread and closely monitoring market order arrivals. Consequently, trade duration and quote

duration are relevant measures for speed of information flow and quote revision, respectively. Our model includes them as two factors influencing liquidity changes and assumes that the trade duration factor is strictly exogenous. The joint modeling of trade duration and corresponding quote duration is challenging due to the fact that they are not synchronic by nature. To circumvent this problem, Engle and Lunde (2003) propose a bivariate point process. As a general conclusion, they found that information flow variables, such as trade duration, large volume of trade and spread, predict more rapid quote revision. We first apply their bivariate model to our Xetra dataset and then extend the analysis to the impact of both trade and quote durations on ex-ante liquidity changes.

Second, by applying a decomposition model, we perform a much finer analysis of ex-ante liquidity changes and take advantage of econometric modeling by attempting to capture a more general and realistic LOB trading pattern which is much more complete than that characterized by structural models. Specifically, following Engle and Lunde (2003) and Rydberg and Shephard (2003), we use different factors to model the dynamics of liquidity changes, including trade duration, quote duration, activity, direction, and size. Our objective is to identify the possible determinants of each factor from a wide range of variables. Our empirical findings will not only provide support for existing theoretical models but will also offer a guidance for new theoretical models in market microstructure.

Third, regarding the explanatory variables, in addition to the lagged dependent variables, we also include various exogenous variables in the different factor equations. In existing literature, most papers take one variable as the explanatory variable and suppose that this variable can summarize all the trade information. Instead, our variables are volume-related, duration-related and trade imbalance related. Among these variables, we also distinguish between the short-run and long-run<sup>4</sup> variables to reflect their time dimension. Our results suggest that most trade-related variables have an impact on the activity factor except the long-run trade imbalance variable. Transaction quantity has very different effects on liquidity measure based on different volumes, and there is a leverage effect such that the magnitude of liquidity decrease is higher than that of increase.

The rest of the paper is organized as follows: Section 2 describes the Xetra trading system and the ex-ante liquidity measure Xetra Liquidity Measure (XLM) we model in this study. Section 3 presents the decomposition model and the exogenous variables we use for explaining the dynamics of each factor. Section 4 applies our econometric model to the data for the selected stocks and reports the estimation results. Section 5 concludes and provides possible new research directions.

---

<sup>4</sup>Short-run variables are variables at a given timepoint, whereas the long-run variables are variables that summarize the information over an interval.

## 2 Xetra trading system and Xetra Liquidity Measure (XLM)

### 2.1 Xetra trading system

Electronic trading systems have been adopted by many stock exchanges during the last two decades. The data used in this study are from the trading system Xetra, which is operated by Deutsche Börse at Frankfurt Stock Exchange (FSE) and has a similar structure to the Integrated Single Book of NASDAQ and Super Dot of NYSE. Xetra trading system realizes more than 90% of the total transactions at German exchanges. Since September 20, 1999, trading hours have been from 9h00 to 17h30 CET (Central European Time). However, during pre- and post-trading hours, entry, revision and cancellation are still permitted.

There are two types of trading mechanism during normal trading hours: the call auction and the continuous auction. A call auction can be organized once, or several times during the trading day in which the clearance price is determined by the state of LOB and remains as the open price for the following continuous auction. During each call auction, market participants can submit both round-lot and odd-lot orders, and both the start and end time for a call auction are randomly chosen by a computer to avoid scheduled trading. Between the call auctions, the market is organized as a continuous auction where traders can only submit round-lot-sized limit orders or market orders.

For highly liquid stocks, there are no dedicated market makers during continuous trading. As a result, all liquidity comes from limit orders in the LOB. The Xetra trading system imposes a Price Time Priority condition, where the electronic trading system places the incoming order after checking the price and timestamps of all available limit orders in the LOB. Our database includes 20 levels of LOB information<sup>5</sup> meaning that, by monitoring the LOB, any registered member can evaluate the liquidity supply dynamic and potential price impact caused by a market order. However, there is no information on the identities of market participants.

The reconstruction of the LOB is mainly based on two main types of data streams: delta and snapshot. The delta tracks all the possible updates in the LOB such as entry, revision, cancellation and expiration, whereas the snapshot gives an overview of the state of the LOB and is sent after a constant time interval for a given stock. Xetra original data with delta and snapshot messages are first processed using the software XetraParser developed by Bilodeau (2013) in order to make Deutsche Börse Xetra raw data usable for academic and professional analysts. XetraParser reconstructs the real-time order book sequence including all the information for both auctions and continuous trading by implementing the Xetra trading protocol and Enhanced Broadcast. We then put the raw LOB information in order and in a readable format for each update time and we retrieve useful and accurate information about the state of the LOB and

---

<sup>5</sup> The hidden part of an iceberg order is not observable in our dataset.

the precise timestamp for order modifications and transactions during continuous trading. The stocks Metro AG (MEO), Merck (MRK), RWE AG (RWE) and ThyssenKrupp AG (TKA) that we choose for this study are blue chip stocks from the DAX30 index. The selected stocks have different levels of market capitalization and are in different markets. Metro AG, with market capitalization of 9.8 billion Euros, operates retail stores, supermarkets and hypermarkets on-line and off-line. Merck is the world's oldest operating chemical and pharmaceutical company with a market capitalization of 4 billion Euros in 2010. RWE generates and distributes electricity to various customers including municipal, industrial, commercial and residential customers. The company produces natural gas and oil, mines coal and delivers and distributes gas. In 2010, its market capitalization was around 15 billion Euros. ThyssenKrupp AG manufactures iron and steel industrial components with market capitalization of 11.9 billion Euros.

## 2.2 (Il)liquidity measure in the Xetra Trading System

As mentioned above, liquidity is the central quality criterion for the efficiency of marketplaces in electronic securities trading. We use the following definition of liquidity: the ability to convert the desired quantity of a financial asset into cash quickly and with little impact on the market price (Demsetz (1968) ; Black (1971); Glosten and Harris (1988)). The Xetra trading system defines its own ex-ante (il)liquidity measure Xetra Liquidity Measure (XLM) as follows<sup>6</sup>:

$$XLM^q = \frac{P_{net,buy}^q - P_{net,sell}^q}{P_{mid}} \times 10000 \quad (1)$$

$$P_{net,buy}^q = \frac{\sum_{k=1}^{K-1} P_{k,i} \cdot v_{k,i} + P_{K,i} \cdot v_{K,i}}{v} \quad \text{and} \quad v_{K,i} = v - \sum_{k=1}^{K-1} v_{k,i}$$

where  $q$  is the potential size in Euro, the conventional value for  $q$  in Xetra trading system is 25 000 Euros<sup>7</sup>.  $P_{net,buy}^q$  is the average price when a buy market order of  $q$  euros arrives and  $P_{net,sell}^q$  relates to the average price for a sell market order of  $q$  euros.  $v$  is the total volume bought by the market order of  $q$  euros.  $P_{k,i}$  and  $v_{k,i}$  are the  $k$ th level ask price and volume available, respectively.  $v_{K,i}$  is the quantity left after  $K - 1$  levels are completely consumed by the market order of  $q$  euros.  $P_{net,sell}^q$  is computed in a similar way.  $P_{mid}$  is the mid-quote of bid-ask spread. XLM measures the relative potential round-trip impact when buying and liquidating a volume of  $q$  euros at the same time. Intuitively, it is also the cost in basis points for an immediate

<sup>6</sup>More information is available on the official website of Xetra.

<sup>7</sup>In this study,  $q$  takes values of 5 000, 10 000, 15 000 and 25 000 Euros.

demand for liquidity from buy and sell market orders. For example, an XLM of 10 and a market order volume of 25 000 Euros means that the market impact of buying and selling is 25 Euros. By choosing different volume  $q$ , we can identify the ex-ante liquidity embedded in the open LOB. The dynamic of  $XLM^{5000}$  reflects the evolution of ex-ante liquidity embedded in lower levels of LOB, whereas  $XLM^{25000}$  captures the ex-ante liquidity embedded in both lower and higher levels of LOB. The previous market microstructure literature that considers the quantity available in LOB includes Irvine et al. (2000), Domowitz et al. (2005), and Coppejans et al. (2004), among others.

XLM is based on  $P_{mid}$  and the difference between  $P_{net,buy}^q$  and  $P_{net,sell}^q$ . Theoretically, there are infinite combinations of  $P_{net,buy}^q$  and  $P_{net,sell}^q$  for the same difference. That is, the illiquidity may come from either side or both sides of LOB. However, our study focuses on stocks' global liquidity and considers the lack of depth in either side as an illiquid situation. For the buy side or sell side investors, the corresponding one-side XLM can be defined and computed in a similar fashion.

The XLM measure can be used for several ends: first, it can help decision making in security selection when constructing a portfolio. Among the stocks with same correlation with market portfolio, a small XLM stock will decrease the trading cost and in the end provide a higher net return. Second, XLM can also be used for comparison purposes. For instance, a cross listing stock may have different liquidity features in different markets. By using XLM, one can quantify this difference by choosing a given volume.

## 3 Methodology

### 3.1 Model

In our dataset, there are three variables to model: trade, quote and ex-ante liquidity changes. Following Engle and Lunde (2003), we consider trade and quote as a bivariate point process. Based on the timestamps of these point processes, we can define two types of duration: trade duration and quote duration which constitute a bivariate duration process. However, due to the no-synchronization problem, we further assume that the trade times are the initiators for both the following trade and the next quote update. Consequently, trade durations and quote durations with the same index share the same original timestamp. The economic intuition behind the assumption is that the limit order traders in open LOB update their quotes by observing the transactions. After each transaction, we compute the quote duration based on the very last transaction.

As mentioned by Engle and Lunde (2003), by taking the transaction times as the origin of each pair of durations, two possible situations may occur for quote duration: an uncensored



observation or a censored observation. The uncensored duration occurs when the quote update is before the next trade arrival and the censored duration happens when the following trade arrives before the quote update. We denote  $x_i$  and  $y_i$  as the trade duration and quote duration, respectively, and further define the observed quote duration  $\tilde{y}_i = (1 - d_i) \cdot y_i + d_i x_i$ , where  $d_i = I_{\{y_i > x_i\}}$ .

Apart from the time dimension, for our purposes, we have to model the liquidity dimension. We use the above mentioned XLM as the liquidity measure. In the uncensored situation, we take the average of the measure within the first quote update timestamp and the following trade timestamp. We propose that its evolution could be written as:

$$XLM_i^q = XLM_0^q + \sum_{k=1}^i Z_k \quad (2)$$

where  $Z_k$  is defined as the  $k$ th rounded signed change for  $XLM^q$ .

We define  $p$  as the joint density for trade duration, quote duration and  $XLM^q$  changes. We propose the following decomposition for this joint density of  $k$ th mark:

$$\begin{aligned} & p(x_k, \tilde{y}_k, z_k^q | F_{k-1}; \omega) \\ &= g(x_k | F_{k-1}; \omega_1) \cdot f^{Dur}(\tilde{y}_k | x_k, F_{k-1}; \omega_2) \cdot f^A(A_k | x_k, \tilde{y}_k, F_{k-1}; \omega_3) \cdot \\ & f^D(D_k | x_k, \tilde{y}_k, A_k, F_{k-1}; \omega_4) \cdot f^S(S_k | x_k, \tilde{y}_k, A_k, D_k, F_{k-1}; \omega_5) \end{aligned}$$

We define  $x_k$  and  $\tilde{y}_k$  as exogenous factors, which relate to the trade duration and observed quote duration, for  $XLM^q$  changes. Conditional on information set  $F_{k-1}$  and two durations,  $A_k$  takes on value 0 or 1 indicating if there is a change on  $k$ th  $XLM^q$ . Conditional on  $A_k = 1$ ,  $D_k$  relates to the direction of the  $XLM^q$  change by taking on the value  $-1$  and  $+1$ . Finally, given the information set  $(F_{k-1}, A_k = 1, Dir_k)$ ,  $S_k$  takes on positive integers and indicates the size of the change.  $\omega$  relates to the parameter set including  $\omega_1, \omega_2, \omega_3, \omega_4$  and  $\omega_5$  which are the parameters for factors of trade duration, observed quote duration, activity, direction and size.

We adopt the Log-ACD model originally introduced by Bauwens and Giot (2000) in modeling the irregularly spaced trade durations which represents a main characteristic of high frequency data:

$$\frac{x_k}{\psi_k} = \varepsilon_k, \quad \psi_k = \exp \left( \omega + \sum_{j=1}^p \alpha_j \varepsilon_{k-j} + \sum_{j=1}^l \beta_j \ln \psi_{k-j} + \Psi' W_{k-1} \right) \quad (3)$$

where  $\psi_k = E(x_k | F_{k-1})$  and  $\varepsilon_k$  is a i.i.d random variable following the generalized gamma distribution with unit expectation. The process of duration is composed of a sequence of deseasonalized durations.  $W_{k-1}$  is a vector of exogenous variables available at  $k - 1$  which includes

trade-related variables, quote-related variables and dummy variables to capture the intraday seasonality (hereafter, all exogenous vectors will include dummy variables to capture the intraday seasonality). We use the Log-ACD as it is more flexible in modeling and the positivity constraint on durations is always respected<sup>8</sup>.

We adopt a similar structure for observed quote duration, that is

$$\frac{\tilde{y}_k}{\phi_k} = \epsilon_k, \quad \phi_k = \exp \left( \mu + \sum_{j=1}^p \rho_j \epsilon_{k-j} + \rho_{j+1} \epsilon_{k-1} d_{k-1} + \sum_{j=1}^l \delta_j \ln \phi_{k-j} + \Phi' V_{k-1} \right) \quad (4)$$

where  $\phi_k = E(\tilde{y}_k | F_{k-1})$  and  $\epsilon_k$  is supposed to be i.i.d exponential distributed and the error distribution is supported by the estimation convergence. In the equation, we add a term with a censored dummy variable to capture the impact of censored observation and the vector of exogenous variables  $V_{k-1}$  which may have some common variables with  $W_{k-1}$ . We show in the following section that as quote durations are conditional on transactions, one possible exogenous variable could be the expected trade duration available at tick  $k$ .

Regarding the liquidity dimension, we decompose the  $XLM^q$  change into three factors: Activity, Direction and Size. The advantage of the decomposition approach is that we can use simple and interpretable factors to model the different facets of liquidity. To this end, the first factor, Activity, is a bivariate variable that takes values of 0 or 1 to indicate whether there is a change in the  $XLM^q$  measure. To model the activity factor, we use the auto-logistic model (Cox et al., (1981)). As the log-likelihood function of the auto-logistic is concave, the numerical optimization can be easily and reliably realized. However, the high-frequency data often performs a slow decay for longer lags in an autoregressive structure. Thus there is a trade-off between bias and variance, i.e., inference with too few parameters may be biased, while that with too many parameters may cause precision and identification problems. To solve this, we adopt another structure called the GLARMA (Generalized Linear Autoregressive Moving Average) binary model which is a generalized structure of auto-logistic structure allowing for moving average-type behavior (Shephard (1995)). The auto-logistic model for activity is defined as:

$$f(A_k = 1 | F_{k-1}, x_k, \tilde{y}_k) = p(\theta_k^A), \text{ where } p(\theta_k^A) = \frac{\exp(\theta_k^A)}{1 + \exp(\theta_k^A)} \quad (5)$$

$$\text{and } \theta_k^A = \Pi'_A M_{k-1}^A + g_k^A, \quad g_k^A = \sum_{j=1}^p \gamma_j^A g_{k-j}^A + \sum_{j=1}^l \lambda_j^A A_{k-j}$$

---

<sup>8</sup> The ACD (Autoregression Conditional Duration) model is initialized by Engle and Russell (1998) and widely used for duration modeling.

Consequently,  $f(A_k = 0 | F_{k-1}, x_k, \tilde{y}_k) = \frac{1}{1 + \exp(\theta_k^A)}$

where  $M_{k-1}^A$  is the vector of exogenous variables for the activity factor known at  $k - 1$ . In this logistic modeling, the parameter  $\theta_k^A$  is time-varying and depends on both its own lag variables, such as lags of  $g_k$  and  $A_k$ , and some exogenous variables. The model will be validated by applying the Ljung-Box test on the standardized errors defined by :

$$u_k^A = \frac{A_k - p(\theta_k^A)}{\sqrt{p(\theta_k^A)(1 - p(\theta_k^A))}} \quad (6)$$

which should be uncorrelated with zero mean and unit variance.

In a similar way, the Direction of change of the liquidity measure conditional on the activity factor is specified by another binary process on 1 or  $-1$  (positive indicates that more liquidity cost should be paid when trading volume  $q$  and negative is related to less liquidity cost situation) and is estimated by another auto-logistic model:

$$f(D_k = 1 | F_{k-1}, x_k, \tilde{y}_k, A_k = 1) = p(\theta_k^D), \text{ where } p(\theta_k^D) = \frac{\exp(\theta_k^D)}{1 + \exp(\theta_k^D)} \quad (7)$$

$$\text{and } \theta_k^D = \Pi_D' M_{k-1}^D + g_k^D, \quad g_k^D = \sum_{j=1}^p \gamma_j^D g_{k-j}^D + \sum_{j=1}^l \lambda_j^D D_{k-j}.$$

Consequently,  $f(D_k = -1 | F_{k-1}, x_k, \tilde{y}_k, A_k = 1) = \frac{1}{1 + \exp(\theta_k^D)}$ .

$M_k^D$  is a vector including exogenous variables of subset  $F_{k-1}$  and  $\Pi^D$  is a parameter vector. It should be noted that the vectors  $M^D$  and  $M^A$  might have some identical exogenous variables.

Once the model is estimated, we use the Ljung-Box test to validate its ability to capture the main features of data. The test will be applied to standardized residuals:

$$u_k^D = \frac{D_k - (2p(\theta_k^D) - 1)}{2\sqrt{p(\theta_k^D)(1 - p(\theta_k^D))}} \quad (8)$$

Finally, the last factor is Size which captures the magnitude of the change of XLM. Less than half of observations in the sample have no size change, that is, they stay at their previous level. Thus we will adopt a geometric process for size changes. The choice of geometric distribution is motivated by its simplicity and generality:

$$S_k | F_{k-1}, x_k, \tilde{y}_k, A_k = 1) \sim 1 + g(\lambda_k) \quad (9)$$

$$\lambda_k = \frac{\exp(\theta_k^{Siz})}{1 + \exp(\theta_k^{Siz})},$$

$$\theta_k^{Siz} = \Pi'_{Siz} M_{k-1}^{Siz} + g_k^{Siz} \text{ and } g_k^{Siz} = \sum_{j=1}^p \gamma_j^{Siz} g_{k-j}^{Siz} + \sum_{j=1}^l \lambda_j^{Siz} S_{k-j}$$

where  $M_{k-1}^{Siz}$  is a vector of exogenous variables and  $\Pi_{Siz}$  is the corresponding parameter vector.  $g(\lambda_k)$  indicates the geometric distribution with parameter  $\lambda_k$ <sup>9</sup>. In order to capture the asymmetry between up-move size and down-move size, we add a direction variable in the vector of exogenous variables. In Equation (9), we add one to the geometric distribution since the minimum change is one. We also apply the Ljung-Box statistics to standardized residuals to evaluate the model. Given the conditional distribution of Size, we have

$$E(S_k - 1 | F_{k-1}, x_k, \tilde{y}_k, A_k = 1) = \frac{1 - \lambda_k}{\lambda_k} \quad (10)$$

$$Var(S_k - 1 | F_{k-1}, x_k, \tilde{y}_k, A_k = 1) = \frac{1 - \lambda_k}{(\lambda_k)^2}. \quad (11)$$

Standardized residuals are computed as

$$u_k^{Siz} = \frac{S_k - 1 - E(S_k - 1 | F_{k-1}, x_k, \tilde{y}_k, A_k = 1)}{\sqrt{Var(S_k - 1 | F_{k-1}, x_k, \tilde{y}_k, A_k = 1)}} \quad (12)$$

and an adequate modeling requires  $u_k^{Siz}$  being uncorrelated with zero mean and unit variance.

In summary, for the estimation process, we can separately estimate each factor by using the Maximum Likelihood approach. The BIC criteria will be applied for the model selection, especially for the choice of number of lags. Moreover, in order to test if the model captures the main features of time series data, we perform a portmanteau test in which residuals will be used to calculate the Ljung-Box statistics as a measure of residual dependence.

Then, with the previous specifications, all the observations can be classified into one of the three following categories:

- 1) There is no change in XLM, that is, activity factor  $A_k = 0$  and no direction and size factors.
- 2) Liquidity decreases and the size change is at least one unit. The corresponding factors are :  $A_k = 1$ ,  $D_k = 1$ , and  $S_k = s_k$ ;
- 3) Liquidity increases and size change is at least one unit. The corresponding factors are :  $A_k = 1$ ,  $D_k = -1$ , and  $S_k = s_k$ ;

---

<sup>9</sup>The general probability distribution function is  $f(x = m) = \lambda(1 - \lambda)^m, 0 < \lambda < 1, m = 0, 1, 2, \dots$

The maximum likelihood estimation function is equal to :

$$\begin{aligned}
& L(\omega_1, \omega_2, \omega_3, \omega_4, \omega_5; x_k, \tilde{y}_k, z_k^q) = \\
& = \sum_{k=1}^n \left[ \begin{aligned}
& \log[g(x_k | F_{k-1}; \omega_1)] + \log[f^{Dur}(\tilde{y}_k | x_k, F_{k-1}; \omega_2)] \\
& + I_k(1) \cdot [\log(f^A(A_k | F_{k-1}; \omega_3))] \\
& + I_k(2) \cdot [\log(1 - f^A(A_k | F_{k-1}; \omega_3)) + \log(f^D(D_k | F_{k-1}; \omega_4)) + \log(f^S(S_k | F_{k-1}; \omega_5))] \\
& + I_k(3) \cdot [\log(1 - f^A(A_k | F_{k-1}; \omega_3)) + \log(1 - f^D(D_k | F_{k-1}; \omega_4)) + \log(f^S(S_k | F_{k-1}; \omega_5))]
\end{aligned} \right] \quad (13)
\end{aligned}$$

where  $I_k(1)$ ,  $I_k(2)$ ,  $I_k(3)$  correspond to the indicator function relating to the three categories mentioned above.

To conclude, the advantage of this modeling is that the partition enables us to simplify the modeling and computation task by specifying the suitable econometric models for the marginal densities of trade duration and conditional densities for quote duration and factors such as Activity, Direction and Size. In addition, for different purposes, the model could also be extended to a more or less complicated context by including other factors. In these decomposition models, one of the crucial tasks is to identify the exogenous variables. Apart from the irregularly spaced duration, it is reasonable to test whether the rest of the factors also contain seasonality effects.

### 3.2 Exogenous Variables Set

Given the model defined above, we need to identify the possible exogenous variables, apart from own lags, for each component. In this study, we attempt to find some variables that have economic interpretation. In previous literature, the most widely used variables have been spread, trading volume and price (Hasbrouck (1996), Goodhart and O'Hara (1997), Coughenour and Shastri (1999) and Madhavan (2000)). The intuition is that trading activities and LOB trader behavior are related. For instance, in a volatile trading period, trading volume will increase and trade duration and quote duration will decrease. Consequently, these variations will generate a volatile open LOB.

The first exogenous variable is relative spread change which is computed by the following formula:

$$RelativeSpread_k = 100 \cdot (\ln(ask_k) - \ln(bid_k))$$

Its variation is measured by:

$$DeltaSpread_k = RelativeSpread_k - RelativeSpread_{k-1}$$

where  $ask$  and  $bid$  are the best sell price and buy price available in open LOB. The advantage of relative spread is that it is dimensionless and can be used to directly compare different stocks.

As the relative spread captures quasi-instantaneous information and might be noisy, another spread-related variable is the average relative spread over the ten most recent observations:

$$AveSpread_k = \frac{1}{10} \sum_{i=1}^{10} RelativeSpread_{k-i}$$

Regarding the volume dimension, the first exogenous variable is the square root of the volume (number of shares),  $SquareRoot(vol)$ , that initiates the current trade. There are two reasons for the use of square root, one is to weigh down the large trade volume, and the second is that the price impact proves to be a concave function of market order size (Hasbrouck (1991)). If the volume that initiates the current trade is large, we expect a volatile situation and, by consequence, the trade duration and quote duration are likely to be short.

The second volume-related variable should capture the imbalance of the signed trade. To this end, we adopt the depth measure proposed by Engle and Lange (2001) which is defined as follows:

$$Abs(sign.vol)_k = \left| \sum_{i=1}^{10} sign_{k-i} volume_{k-i} \right|$$

where the  $sign_{k-i}$  and  $volume_{k-i}$  are the trade sign and trade volume for the  $(k-i)th$  trade. The trades are classified into buy-initiated and sell-initiated according to the rule of Lee and Ready (1991). Intuitively, when the depth measure increases, it indicates that the trades are imbalanced and the market is dominated by one-sided pressure.

The third dimension is duration. We define two sorts of duration, back-quote duration and quote-quote duration, that are different from trade duration and quote duration. The former is used to consider the duration between the first update of LOB after the previous trade and the following trade which contains quote information. It should be noted that the back-quote duration could be zero due to the fact that the quote duration might be censored when the trade occurs before the update of open LOB. The way by which the data is sampled ignores some quotes when there is more than one update between two trades. It might not be a concern when 75% of the quotes are preserved as in Engle and Lunde (2003). However, in a market where the open LOB is more active as in the Xetra trading system, ignoring the quote activity may be a concern. In fact, only around 20% of the quotes are preserved in our dataset.

The quote-quote duration variable considers the duration for which there is no change of XLM. As a result, it will be used only in explaining the components such as direction and size when the liquidity measure changes.

The above variables will be all (or partially) included in the exogenous variable vectors for different components. In addition, we also put time-of-day dummy variables in the vectors to remove the seasonality, a stylized fact in high-frequency data. There exist several techniques to this end. In our study, we use eight time-of-day dummy variables to capture the intraday seasonality for each component for which we desire modeling: one for the first half hour after the market opens and then one for each hour of the trading day until the market closes.

Table 1 presents the summary statistics of durations and exogenous variables for the first week of July 2010. The number of trades for all stocks ranges from 8 256 to 19 488. The trade frequency is also confirmed by the corresponding average trade durations. That is, stocks with larger number

of trades correspond to shorter trade durations. Regarding the average trade volumes, there is a big difference: it varies from 5,180 to 15,472 Euros, meaning that the selected stocks have different levels of liquidity. The average quote durations are relatively small and there is evidence that the dynamics in the open LOB are more active than that of trades. Considering other exogenous variables, it is natural to see that the averages of *DeltaSpread* and  $\Delta XLM$  are close to zero. As *AveSpread* is dimensionless, we can consider this variable as an indicator of transaction costs. All stocks have an average of *AveSpread* around 0.05%, meaning that the average spread remains stable across the stocks. The average trade imbalance variable *Abs(sign.vol)* varies from 944 to 3 616 shares, indicating the existence of different trading patterns across the stocks.

[Insert Table 1 here]

## 4 Estimation and results on liquidity

The data covers the first week of July 2010 and the last week of June 2011 for stocks RWE, MEO, MRK and TKA. We present and compare the estimation results for  $XLM^q$  with potential size  $q$  of 5 000 and 25 000 Euros during the first week of July 2010<sup>10</sup>.  $XLM^{5\,000}$  and  $XLM^{25\,000}$  represent ex-ante liquidities embedded in lower and higher levels of LOB, respectively. By taking different potential sizes, we can examine the dynamics of lower- and higher- levels of LOB with various exogenous variables. The log-likelihood is maximized by the quadratic hill climbing method and the maximization program is run with Matlab v7.6.0 with Optimization Toolbox.

### 4.1 Temporal Factors

The estimation results of trade durations and quote durations are presented in Table 2 and Table 4. Panel A presents the estimation results for lagged dependent variables, and Panel B for exogenous variables. A more detailed analysis can be found in the Appendix. Table 3 and Table 5 show the statistics related to estimations.

#### 4.1.1 Trade Duration Factor

In the modeling, we decompose the explanatory variables of trade durations into two groups: lagged dependent variables and exogenous variables. More specifically, the lagged dependent variables are used to capture the degree of persistence in the trade durations. One part of the exogenous variables will test the effect of different variables on trade durations, and the other part, with dummy variables, helps to remove seasonality.

<sup>10</sup>Results for potential size of 5 000 Euros for the same week are presented in the appendix. More results on other potential sizes and the other week are available upon request.

The overall results on trade durations, presented in Panel B of Table 2, are stable across stocks over the two sample periods and provide some new empirical evidence about trade duration dynamics. First, the sums of the coefficients for the first two lagged dependent variables are around 0.9, suggesting that trade durations are highly persistent. Second, the coefficients for *DeltaSpread* and *AveSpread* are all positive and significant for the 4 stocks, showing that when the liquidity decreases, traders will slow down their trading intensities. Moreover, the coefficient for the short-term variable *SquareRoot(vol)* is negative and significant for the 4 stocks, indicating that large trades will increase trading intensity. However, the coefficient for the long-term variable *Abs(sign.vol)* is positive and significant only for RWE and TKA. This means that, for these two stocks, when trade imbalance increases, the trading activity will slow down. Third, trading durations do perform an intraday pattern.

[Insert Table 2 here]

The model is validated by the Ljung-Box statistics. Table 3 presents the results at different lags for trade durations and the standardized residuals of trade durations. The left side of the table shows that there is a high persistence in autocorrelation of the trade duration dynamics. The right side of the table presents evidence that the Log-ACD model is capable of removing this autocorrelation feature in the trade durations since the Ljung-Box statistics have been reduced dramatically and, in most of cases, the hypothesis of no autocorrelation cannot be rejected.

[Insert Table 3 here]

#### 4.1.2 Quote Duration Factor

Similar to the trade duration equation, we also decompose explanatory variables of the quote duration equation into two parts: lagged dependent variables and exogenous variables. We include more exogenous variables in the quote duration equation than in the trade duration equation since we assume that the trade durations are exogenous and can explain the quote duration dynamics. More specifically, the exogenous variables we use in explaining the dynamics of quote durations are: trade-duration-related variables, censored effect variable, *DeltaSpread*, *AveSpread*, *SquareRoot(vol)*, *Abs(sign.vol)*, *BackQuote duration* and  $\Delta XLM$ . Panel B of Table 4 presents the corresponding estimate results.

The effect of exogenous variables on the quote duration can be briefly summarized as follows: first, spread-related variables, such as *DeltaSpread* and *AveSpread*, and the volume-related variable *SquareRoot(vol)* have negative impact on the quote durations. This suggests that when the spread is large and the trading volume is high, LOB traders speed up their revisions. Second, the coefficient for the trade imbalance variable *Abs(sign.vol)* is positive for MEO, RWE and TKA, and is negative for MRK. The positive coefficient of *Abs(sign.vol)* is not very intuitive if



we consider trade imbalance as a result of information asymmetry. One explanation might be that the positive effect is evidence of high-frequency algorithmic trading. The algorithmic (probably not human) traders may benefit from providing liquidity by creating this trade imbalance. For example, if an algorithmic trader wants to buy a given quantity of stocks, on the ask side he can place limit orders in different price levels with different quantities. As a result, the pressure on the sell side increases and then more trades occur on the bid side. In this case, there is no fear related to informed traders, and so quote revision can be postponed.

[Insert Table 4 here]

In Table 5, we find that the Ljung-Box statistics have been largely reduced and, in most of cases, the hypothesis of no autocorrelation cannot be rejected. It should be noted that, despite the large number of lagged variables, TKA still has a relatively high autocorrelation in the standardized residuals.

[Insert Table 5 here]

## 4.2 Liquidity Activity Factor

Up to now, we have analyzed the dynamics of trade durations and quote durations. Conditional on these temporal variables, we can make further analysis on other dimensions of liquidity. As we mentioned above, we decompose the change of the liquidity measure XLM into three components: Activity, Direction and Size. Similar to the time dimension, we also put two groups of explanatory variables into the activity equation: lagged dependent variables and exogenous variables. The  $M^A$  vector includes expected trade duration, expected quote duration, *DeltaSpread*, *AveSpread*, *SquareRoot(vol)*, *Abs(sign.vol)*, *BackQuote duration*,  $\Delta XLM$  and dummy variables.

Table 6 and Table A.5 report the estimated results of the liquidity activity equation for  $XLM^{25000}$  and  $XLM^{5000}$ . The activity process is a binary process in which the value 1 means a change in the liquidity measure XLM. To model the activity factor, we adopt the GLARMA structure introduced by Rydberg and Shephard (2003). In addition to the GLARMA structure, we also include exogenous variables. Panel A of Table 6 presents the estimated results of the GLARMA part. For all stocks, we use a lag set of (1,2) to capture the autocorrelation of the activity factor. Consistent with previous literature, the coefficients of the “GLAR” part are positive and significant, ranging from 0.66 to 0.95. The result suggests a high persistence in autocorrelation for the activity factor. More specifically, there is a cluster effect in activity, namely, the change of liquidity is more likely to be followed by another change.

[Insert Table 6 here]

In this study, we are also interested in the effect of exogenous variables on the dynamics of liquidity. Panel B of Table 6 shows the estimated results for these exogenous variables. For the time dimension variables, expected trade duration and expected quote duration do not have the same effect on the probability of liquidity change. In particular, a longer expected trade duration increases the probability of liquidity change, whereas a longer quote duration decreases this probability. The same effects are also found for the activity factor of  $XLM^{5000}$ . In the tick-by-tick trading framework, as found by Dionne et al. (2009) and Dionne et al. (2015), a longer trade duration will have a positive impact on price volatility. As a result, a longer trade duration will increase the probability of liquidity change. On the other hand, quote duration measures the quote intensity. A longer quote duration means a less active open LOB. Because of this, the quote is likely to be unchanged.

Regarding spread-related variables, both *DeltaSpread* and *AveSpread* have a positive effect on the probability of liquidity change. These two measures are related to liquidity itself. When liquidity decreases, LOB traders are more prudential in their quotes, and are more likely to update their quotes. Therefore, the probability of liquidity change increases. Recall that, in the trade duration equation, *DeltaSpread* and *AveSpread* also have a positive effect on trade duration. We find that different types of traders have different ways to protect themselves. More concretely, when the market is less liquid, market order traders become less active and LOB traders become more active and careful. It should be noted that the activity factor only tells us whether the liquidity changes or not; there is no information on the direction and magnitude of change.

Regarding volume-related variables, *SquareRoot(vol)* and *Abs(sign.vol)* affect the probability of liquidity change in different ways. The coefficients of *SquareRoot(vol)* are both significant at a 5% level for  $XLM^{5000}$  and  $XLM^{25000}$  changes. However, the signs of the coefficients are opposite (positive for  $XLM^{25000}$  changes and negative for  $XLM^{5000}$  changes). The results suggest that the large trades are likely to increase the probability of high level liquidity changes. As mentioned before, large trades are likely to be informative. Under this circumstance, the LOB traders are more likely to review their high-level quotes and then the resulted high-level liquidity changes. On the other side, low-level liquidity is likely to be insensitive to high trade volumes. One possible explanation is that low-level LOB is competitive and features a quick resilience. As in the trade duration equation and quote duration equation, the effect of *Abs(sign.vol)* is again not clear. Only one stock has a significant coefficient. The results suggest that the probability of liquidity change hardly depends on the long-run trade imbalance. Possible explanations are that: First, *Abs(sign.vol)* captures information on the last ten trades and the activity component has a very short memory. Second, algorithm traders might provide liquidity by “creating” this trade imbalance.

Another time dimension variable, *BackQuote duration*, also has a positive impact on the probability of liquidity change. This is in line with the estimate results for trade duration. As we can

see from Panel B of Table 6, the effect of expected trade duration is higher than that of expected quote duration. Therefore, the longer *BackQuote duration* implies a more volatile market and liquidity is likely to be updated. For the liquidity measures  $XLM^{5000}$  and  $XLM^{25000}$ , as expected, the coefficient is positive and significant at a 5% level. This means that when a stock is less liquid, there is more chance that LOB traders review their quote and then the liquidity of this stock changes. Concerning the dummy variables, we find that there is a seasonality pattern only for some stocks in certain time periods; most of the periods do not exhibit a seasonality pattern. The results are shown in Panel C of Table 6.

Similar to trade durations and quote durations, we use the Ljung-Box statistics to validate the model. The statistics for the activity factor and the corresponding standardized residuals are reported in the left side and right side of Table 7, respectively. On the one hand, the activity factor is highly autocorrelated. At five-lag level, the Ljung-Box statistics range from 358 to 1127 for all stocks and the hypothesis of no autocorrelation is rejected at any confidence level. On the other hand, we find that the model which includes the GLARMA part and exogenous variables can capture this autocorrelation feature very well. On the right side of the table, all statistics have been reduced to less than the critical values except one for MRK.

[Insert Table 7 here]

### 4.3 Liquidity Direction Factor

Another component of liquidity measure is direction which is also a binary process: the value 1 means a decrease of liquidity and -1 means an increase of liquidity. Conditional on the activity factor, the direction factor gives more information about the change of XLM. To capture the dynamics of direction, we use a similar variables set as for the other equations. Three stocks take the GLARMA(2,1) structure and one stock takes the GLARMA(3,1). Panel A of Table 8 and Table A.5 illustrate the estimation results for the GLARMA structure. Interestingly, the sums of the “MA” part are all negative and smaller than -0.5, indicating a mean-reverting feature. That is, the increase of liquidity is likely to be followed by a decrease of liquidity, and vice versa.

[Insert Table 8 here]

The exogenous variables we include in the direction equation are: *QuoteQuote duration*, *DeltaSpread*, *AveSpread*, *SquareRoot(vol)*, *Abs(sign.vol)*, *BackQuote duration* and  $\Delta XLM$ . Panel B of Table 8 and Table A.5 present the estimated results of these exogenous variables. In the direction equation, the new variable *QuoteQuote duration* is defined as the duration between two liquidity measure changes. As the direction component is observed only when the activity factor is equal to one, it is more reasonable to use a temporal variable to capture this time interval. Based on Panel B of Table 8 for  $XLM^{25000}$ , we find that three stocks have a positive coefficient for *QuoteQuote duration* and only one is significant. It appears that *QuoteQuote duration* is

not a relevant variable in predicting the direction factor. The same results are also found for  $XLM^{5000}$ .

Considering the spread related variables, a high *DeltaSpread* increases the probability of liquidity decrease when there is a liquidity change. Intuitively, when the spread increases, this means the LOB traders keep away from mid-quotes, and so XLM is likely to increase. Compared to *DeltaSpread*, the *AveSpread* has the opposite effect on the direction factor. This suggests that the permanent increase of the spread will indeed increase the probability of liquidity increase. It seems evident that when the traders have to pay a higher liquidity premium, the LOB traders (i.e, liquidity providers) are willing to provide liquidity.

Regarding the volume-related variables, the coefficient of *SquareRoot(vol)* is positive for  $XLM^{25000}$  meaning that the current large trades predict a less liquid situation for a large trade, which is consistent with previous literature. Intuitively, the large trade is likely to be informed trade which will create volatility in the market. Accordingly, when LOB traders are likely to keep away from mid-quotes, and so the liquidity will decrease. However, as shown in Table A.7, *SquareRoot(vol)* has no effect on the low level liquidity direction factor, suggesting that low level liquidity is less sensitive to trade volumes. The trade imbalance variable *Abs(sign.vol)* has no significant effect on the direction factor for both  $XLM^{5000}$  and  $XLM^{25000}$  changes. This confirms the results for the activity component. There is no clear impact of *Abs(sign.vol)* on liquidity change, therefore the effect of *Abs(sign.vol)* on direction is not clear. As we focus on the dynamics of liquidity measure XLM, the reason to the lack of information content of trade imbalance is out of our scope. However, it might be interesting for trading strategy analysis or algorithm trading analysis to continue exploring this aspect.

Another temporal variable, *BackQuote duration*, has a positive impact on the probability of liquidity increase. This suggests that when trades become less active, the LOB traders are likely to review their quotes and incite trades by providing more liquidity. Regarding liquidity measure  $XLM^{5000}$  and  $XLM^{25000}$ , an increase of XLM will increase the probability of being liquid. This is the evidence of the mean-reverting in XLM dynamics. As for the other components, we also use dummy variables to capture seasonality in the direction factor. The estimated results show that the coefficients are not significant, meaning that there is not a clear seasonality effect on the direction factor. The results for coefficients of dummy variables are presented in Panel C of Table 8.

The direction factor is observed when the liquidity measure XLM changes. The Ljung-Box statistics for the direction factor and corresponding standardized residuals are reported in Table 9, respectively. Similar to the activity factor, the direction factor is highly autocorrelated. The hypothesis of no autocorrelation is rejected at any confidence level for all stocks. That is, the Ljung-Box statistics for standardized residuals from 5 to 200 lags are not significant at a 5% level.

[Insert Table 9 here]

## 4.4 Liquidity Size Factor

The last factor is the size of liquidity change. Table 10 and Table A.9 report the estimated results for both  $XLM^{25000}$  and  $XLM^{5000}$  changes. Panel A of Table 10 shows the results of the GLARMA part. We find that the number of lags of the GLARMA part ranges from (2,1) to (2,2). Depending on the stock, the effect of lagged value is either positive or negative. As shown in equation 10, a higher  $\lambda_k$  indicates a smaller expectation of size factor. Combined with the estimated results of the direction component, the implication is that XLM does have a mean-reverting feature, however, the magnitude of reverting varies from one stock to another. Similar results are found in Table A.9 for size factor of  $XLM^{5000}$  changes.

[Insert Table 10 here]

The exogenous variables are the same as in the equation for the liquidity direction factor. In addition, we include current direction and lagged direction in the size equation. In total, there are nine exogenous variables in the size equation. Panel B of Table 10 and Table A.9 report the estimated results of exogenous variables for  $XLM^{25000}$  and  $XLM^{5000}$ , respectively. Once we have all results for liquidity activity, direction and size factor, we can closely analyze the entire effect of exogenous variables on every dimension of liquidity changes. Table 12 presents and compares the effects of key exogenous variables on  $XLM^{25000}$  and  $XLM^{5000}$  changes and liquidity measures.

In line with the estimated results for the direction factor, *QuoteQuote duration* does not have a significant effect on all stocks with respect to both  $XLM^{25000}$  and  $XLM^{5000}$  changes. Recall that *QuoteQuote duration* is the duration between two XLM changes. Again, it indicates that the LOB structure has a short memory and does not have the power to predict dynamic of size change.

As shown in Panel B of Table 10, there exists a leverage effect. That is, the negative size factor change (less liquid) is higher than that of positive change (more liquid). Consistent with Rydberg and Shephard (2003), the current direction variable has negative and significant impact on the  $\lambda_k$ , which implies a higher expected liquidity change when liquidity decreases, and confirms the leverage effect for both  $XLM^{25000}$  and  $XLM^{5000}$  changes. However, the lagged direction is less significant than the current one.

The temporal variable *BackQuote duration* has positive and significant effect on  $\lambda_k$  for both  $XLM^{25000}$  and  $XLM^{5000}$  changes, indicating that even though the liquidity provider tries to incite the traders to trade by increasing the liquidity provision (our conclusion from the results of the direction equation), the magnitude of liquidity increase is moderated. On the other side, if the trading intensity increases and then the *BackQuote duration* decreases, the liquidity is likely to decrease with a greater magnitude.

The last exogenous variable  $\Delta XLM$  itself has a negative and significant effect on  $\lambda_k$  for both  $XLM^{25000}$  and  $XLM^{5000}$  changes. This suggests a quick resilience in liquidity. If the market is evaluated previously as less liquid based on XLM, the actual liquidity is prone to increase and the size of this increase is likely to be large. On the other side, if the market is evaluated as liquid by the XLM, the actual liquidity is likely to decrease with a moderated size factor.

We also attempt to test if there is a seasonality in the size factor by using dummy variables. As shown in Panel C of Table 10, similar to activity and direction factors, there is little seasonality feature in size component because most of the coefficients are not significant from zero at a 5% level.

As for the direction factor, the size factor is also observed when the liquidity measure XLM changes. Table 11 reports the Ljung-Box statistics of the size factor and the corresponding standardized residuals. Interestingly, the statistics for the size factor vary significantly across the stocks. For instance, the size factors of RWE and TKA, with Ljung-Box statistics of 1406 and 889 for 5 lags, are relatively high autocorrelated. Given the high Ljung-Box statistics of standardized residuals for stocks RWE and TKA at small lags (from 5 lags to 20), the model might have a misspecification problem on  $\theta_k^{Siz}$  or a mild distributional failure. It should be noted that  $\theta_k^{Siz}$  can be specified in many different ways and the distribution for size factor can differ from geometric distribution. However, a parsimonious and interpretable model is always preferred.

[Insert Table 11 here]

## 4.5 Summary of Estimation

Up to now, we can compare the total effect on liquidity from different exogenous variables. Table 12 summarizes and compares the effect of key exogenous variables on  $XLM^{25000}$  and  $XLM^{5000}$ . The spread-related variable, *DeltaSpread*, has positive impact on  $\lambda_k$ : three stocks are significant for  $XLM^{5000}$  and only one for  $XLM^{25000}$ . It suggests that the size factor of  $XLM^{5000}$  is more sensitive to the changes in the spread than that of  $XLM^{25000}$ . Combined with the results of the direction factor, it shows that when previous spread increases, the actual liquidity is likely to decrease. Furthermore, the size of this decrease in liquidity is expected to be small. However, if the previous *DeltaSpread* is small, the actual liquidity is likely to increase with a higher size factor. There exists an asymmetry in the effect on the size factor of the liquidity changes.

[Insert Table 12 here]

Panel B and C of Table 12 also compare the different effect of the same exogenous variables on  $XLM^{25000}$  and  $XLM^{5000}$  changes. *AveSpread* has a negative significant effect on  $\lambda_k$  for TKA

with respect to  $XLM^{25000}$  changes and a positive effect for RWE and TKA with respect to  $XLM^{5000}$  changes. The implication is that the dynamics of liquidity only has a short memory. Combined with the results of the direction equation, the negative effect of *AveSpread* on  $\lambda_k$  suggests that high *AveSpread* will lead to a more liquid situation for both  $XLM^{25000}$  and  $XLM^{5000}$  changes. However, the magnitude of this increase in liquidity is larger for  $XLM^{25000}$  changes than that of  $XLM^{5000}$  changes. On the other hand, if the *AveSpread* is small, the actual liquidity will decrease and the magnitude of this decrease in liquidity is smaller for  $XLM^{25000}$  changes than that of  $XLM^{5000}$  changes.

Relative to the size factor of  $XLM^{25000}$  changes, *SquareRoot(vol)* is not significant for three stocks but significant for TKA<sup>11</sup>. However, *SquareRoot(vol)* has a significant positive effect on all the stocks with respect to the size factor of  $XLM^{5000}$  changes, confirming high resiliency at lower levels of LOB. Our model also quantifies the effect of trade quantity on liquidity change. As for the effect of volume on  $XLM^{25000}$ , a high trading volume predicts a decrease of liquidity with a high magnitude for TKA. For the other stocks, liquidity change seems limited to one unit, whereas a low trading volume will lead to an increase of liquidity with a smaller magnitude. Interestingly,  $XLM^{5000}$  is less sensitive to high trading volume than  $XLM^{25000}$ . That is, a high trading volume does not increase the liquidity change probability and the magnitude of the liquidity change. However, a small trading volume is likely to increase the liquidity change probability and the magnitude of this change. This is evidence that liquidity providers at a low-level of LOB are more sensitive to the small trading volumes which are usually issued from predetermined trading strategies.

As shown in Panel B and C of Table 12, the effect of the trade imbalance variable *Abs(sign.vol)* is not significant for size factors for both  $XLM^{5000}$  and  $XLM^{25000}$  changes. Moreover, for  $XLM^{5000}$  changes, there is no effect of imbalance on each of the three factors, confirming the lack of informativeness of trade imbalance. Concerning the  $XLM^{25000}$  changes, given the negative effect of *Abs(sign.vol)* on the direction factor, it suggests that if there is an imbalance in the previous trades, liquidity is likely to increase in a moderated fashion.

## 5 Conclusion

After the introduction of the open LOB trading mechanism, trading frequency has become higher than ever before. By consequence, liquidity has become an important issue for active traders, investors and financial institutions. This paper analyzes of the dynamics of ex-ante liquidity

---

<sup>11</sup> Recall that the value of the size component is defined as the size change minus one. Therefore, the nonsignificance means that the size component changes less than one unit.

changes. The ex-ante liquidity measures used in this study are  $XLM^{25000}$  and  $XLM^{5000}$  proposed by the Xetra trading system. Different from an ex-post liquidity measure, the XLM is an ex-ante volume dependent measure. The computation of the measure requires information such as the prices and the corresponding quantity available in the open LOB.

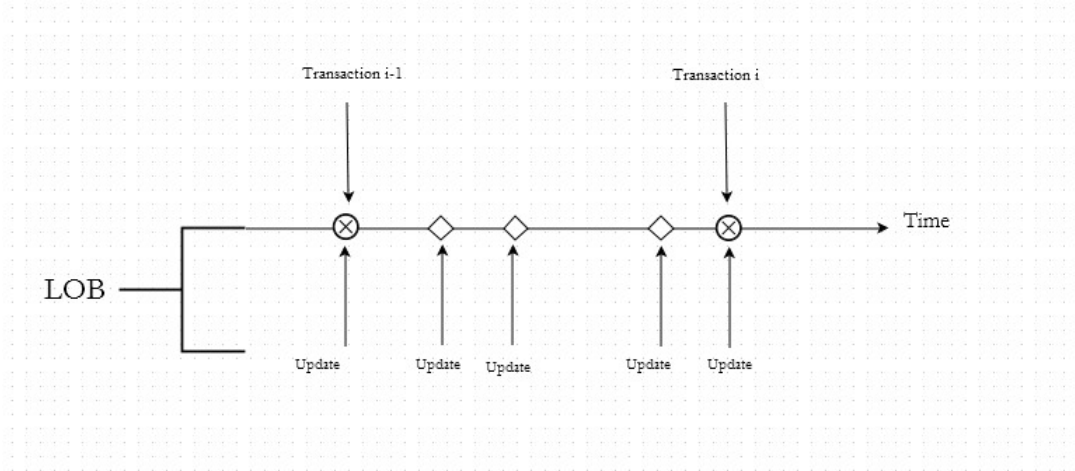
To model the dynamics of the liquidity changes, we adopt the decomposition approach proposed by Rogers and Zane (1998). The liquidity changes have been decomposed into five factors: trade durations, quote durations, activity, direction and size. Trade durations and quote durations are temporal variables and are not synchronized. To manage this problem, we follow Engle and Lunde (2003) by defining the last trade time as the initial quote time in quote duration computation. The other three factors are directly related to the change of XLM itself. Both the activity factor and the direction factor are binary processes taking the value 0 and 1, or -1 and 1, respectively. The size factor captures the magnitude of XLM change. To investigate the dynamics of each component, we apply the relevant econometric models to each factor and put a wide range of trade-related exogenous variables into the models. These include volume-related variables, trade-balance-related variables and temporal variables. The models are validated by the  $t$ -test of the coefficients and the Ljung-Box test on the deduced standardized residuals.

By including different microstructure based variables, we first find that most trade-related variables can influence the dynamic of both trade and quote durations. Moreover, the quote durations are influenced by the dynamics of trade durations, trade-related variables and  $\Delta XLM$ . Second, expected trade durations and quote durations are likely to affect the probability of liquidity change. Most the trade-related variables have an impact on the activity factor except the long-run trade imbalance variable,  $Abs(sign.vol)$ . Third, the temporal variable *QuoteQuote duration* does not seem to have an impact on the direction factor. Fourth, there is a leverage effect in the size factor, that is, the magnitude of liquidity decrease is higher than that of increase. Among the trade-related variables, only spread change has a significant effect on the size factor for  $XLM^{25000}$  and  $XLM^{5000}$  changes. *SquareRoot(vol)* has a very different effect on  $XLM^{25000}$  and  $XLM^{5000}$  changes, suggesting a high resilience at the lower levels of LOB. Fifth, the trade durations and quote durations have an obvious seasonality pattern, whereas the seasonality pattern for other factors is not clear. Overall, we provide evidence that trade durations and quote durations have an impact on ex-ante liquidity. Other exogenous variables that affect ex-ante liquidity include *DeltaSpread*, *AveSpread*, *SquareRoot(vol)*, *BackQuote duration* and  $\Delta XLM$ .

Future research can continue in several directions. Our study focuses on the impact of trade-related variables on liquidity changes. A possible alternative is to investigate how liquidity change co-moves with trades. Another direction is to decompose the liquidity changes in a different order or into different factors to answer other microstructure questions. It could also be interesting to generalize the model from one particular stock to a portfolio. Again, the unsynchronization of the trade durations and quote durations between different stocks is a challenge. It will require a more complicated econometric model and reasonable assumptions.

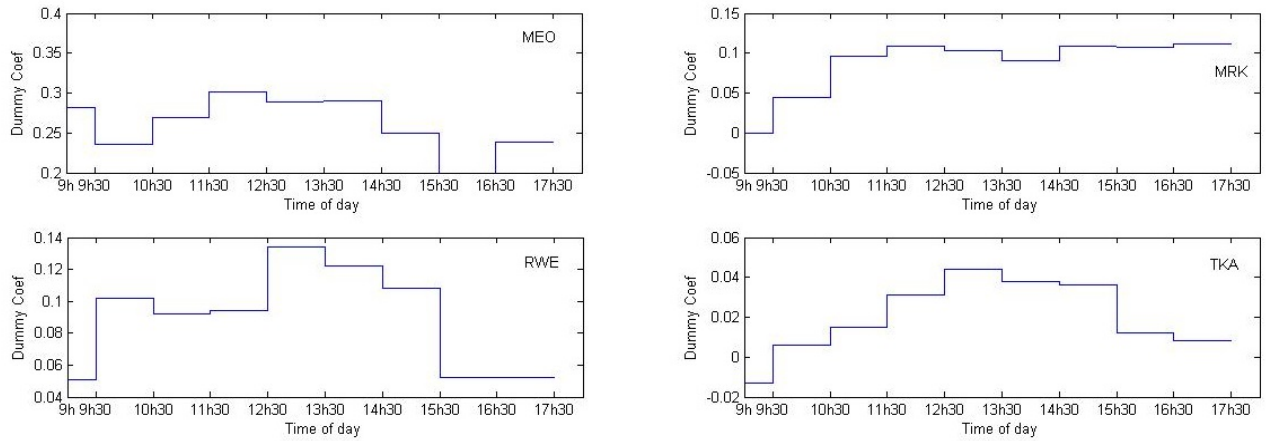


Figure 1: Timestamps for trades and quote update in open LOB



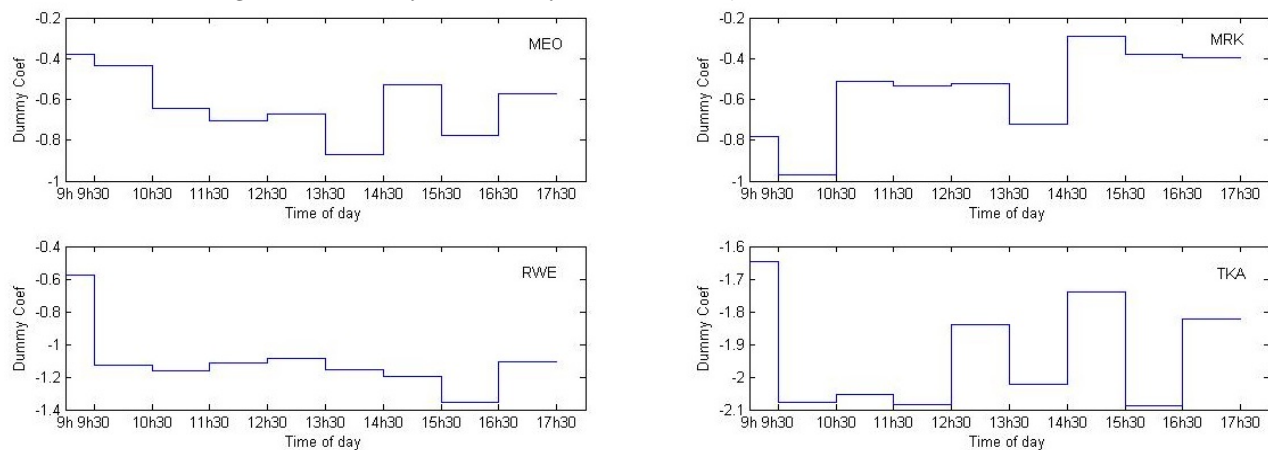
This figure presents the temporal relation between trades and quote updates. The trades and the quote updates are presented by circles and squares, respectively.

Figure 2: Intraday Seasonality Pattern For Trade Durations



This figure presents the intraday patterns of seasonality variables for trade durations for different stocks.

Figure 3: Intraday Seasonality Pattern For Quote Durations



This figure presents the intraday patterns of seasonality variables for quote durations for different stocks.

Table 1: Summary Statistics of Trade Related Variables and Exogenous Variables

Trade Related and Duration Variables						
	Nb.Obs	Average Price	Average Volume	Average Price	Ave.Trade Duration	Ave.Quote Duration
		(euros)	(shares)	(euros)	(seconds)	(seconds)
MEO	10742	43.19	262.00	36.65	14.12	0.12
MRK	8256	59.37	193.74	26.84	18.38	0.14
RWE	16487	54.37	390.84	41.15	9.20	0.16
TKA	19488	21.00	746.05	20.74	7.79	0.14
Average of Exogenous Variables						
	DeltaSpread	AveSpread	SquareRoot(Vol)	Abs(sign.vol)	BackQuote Dur	$\Delta$ XLM
	%	%	(shares)	(euros)	(seconds)	(1/4 bps)
MEO	-1.52E-04	0.049	14.344	1322.493	14.008	-4.75E-03
MRK	-4.77E-05	0.049	12.309	944.487	18.231	-3.75E-03
RWE	-4.01E-05	0.041	17.021	1924.649	9.043	-1.03E-03
TKA	-4.01E-06	0.048	23.538	3616.967	7.653	-4.62E-04

The table reports the summary statistics for trade related variables, durations and exogenous variables. The Nb.Obs is the number of observations in the period of first week of July 2010. The lower part of the table presents the average of the exogenous variables we defined in Section 3.2. Instead of using basic point, we use 1/4 basic point as one unit of liquidity change.

Table 2: Estimated Results For Trade Durations

		MEO			MRK			RWE			TKA		
		Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat
Panel A	ln(dur(lag1))	0.875	0.04	<b>21.904</b>	0.804	0.04	<b>19.958</b>	0.975	0.026	<b>37.655</b>	1.194	0.037	<b>32.168</b>
	ln(dur(lag2))	-0.021	0.034	-0.627	0.126	0.056	<b>2.239</b>	-0.13	0.037	<b>-3.481</b>	-0.306	0.062	<b>-4.969</b>
	ln(dur(lag3))	0.063	0.021	<b>2.958</b>	0.024	0.038	0.629	0.11	0.027	<b>4.045</b>	0.094	0.033	<b>2.848</b>
	x/phi (lag1)	0.081	0.008	<b>10.057</b>	0.06	0.006	<b>9.96</b>	0.119	0.007	<b>17.467</b>	0.13	0.009	<b>14.863</b>
	x/phi (lag2)										-0.059	0.009	<b>-6.81</b>
	gamme1	0.512	0.021	<b>24.364</b>	0.515	0.024	<b>21.552</b>	0.974	0.026	<b>36.793</b>	0.798	0.019	<b>42.379</b>
	gamme2	0.78	0.052	<b>15.077</b>	0.711	0.054	<b>13.17</b>	0.35	0.014	<b>25.582</b>	0.449	0.016	<b>27.562</b>
cons	0.239	0.033	<b>7.147</b>	0.111	0.027	<b>4.128</b>	0.052	0.016	<b>3.282</b>	0.008	0.007	1.087	
Panel B	delta spread	14.848	0.276	<b>53.892</b>	26.521	0.937	<b>28.303</b>	18.192	0.6	<b>30.336</b>	14.341	0.556	<b>25.782</b>
	lev.spread	1.146	0.187	<b>6.119</b>	1.232	0.226	<b>5.444</b>	0.807	0.176	<b>4.597</b>	0.488	0.084	<b>5.777</b>
	square root(vol)	-0.013	0.001	<b>-9.499</b>	-0.008	0.002	<b>-5.075</b>	-0.008	0.001	<b>-12.042</b>	-0.003	0	<b>-9.700</b>
	abs(sign.vol)	4.51E-06	4.94E-06	0.913	3.45E-06	7.05E-06	0.489	4.54E-06	1.64E-06	<b>2.768</b>	1.24E-06	5.30E-07	<b>2.340</b>
	Back_quote_dur	3.90E-04	5.53E-04	0.705	8.17E-04	3.25E-04	<b>2.512</b>	-0.002	0.001	<b>-3.337</b>	-0.002	3.92E-04	<b>-3.971</b>
	$\Delta$ XLM	0.019	0.014	1.404	-0.012	0.013	-0.902	0.019	0.007	<b>2.488</b>	0.001	0.013	0.109
Panel C	Dummy1	0.043	0.024	1.773	-0.112	0.02	<b>-5.492</b>	-0.002	0.011	-0.156	-0.021	0.006	<b>-3.690</b>
	Dummy2	-0.003	0.016	-0.174	-0.066	0.013	<b>-5.191</b>	0.049	0.008	<b>5.925</b>	-0.002	0.004	-0.574
	Dummy3	0.03	0.015	1.991	-0.015	0.012	-1.294	0.04	0.009	<b>4.549</b>	0.007	0.004	1.900
	Dummy4	0.063	0.018	<b>3.542</b>	-0.002	0.012	-0.144	0.041	0.009	<b>4.513</b>	0.023	0.005	<b>4.757</b>
	Dummy5	0.05	0.016	<b>3.165</b>	-0.008	0.014	-0.593	0.082	0.011	<b>7.743</b>	0.037	0.006	<b>6.512</b>
	Dummy6	0.051	0.017	<b>2.939</b>	-0.021	0.012	-1.713	0.069	0.01	<b>7.228</b>	0.03	0.006	<b>4.683</b>
	Dummy7	0.011	0.013	0.885	-0.002	0.012	-0.189	0.056	0.009	<b>6.502</b>	0.028	0.005	<b>5.315</b>
	Dummy8	-0.039	0.011	<b>-3.404</b>	-0.004	0.011	-0.365	-0.001	0.007	-0.088	0.005	0.004	1.202

The table reports the estimated results for trade durations. Panel A presents the results for Log-ACD part, Panel B includes the results for the exogenous variables and Panel C consists of the results for dummy variables. Bold entries indicate the coefficients are significant at a 5% level.

Table 3: Ljung-Box Statistics For Trade Durations and Deduced Standardized Residuals

Ljung-Box Statistics on Trade Durations						Ljung-Box Statistics on Corresponding Standardized Residuals					
Lags	MEO	MRK	RWE	TKA	C_Value	Lags	MEO	MRK	RWE	TKA	C_Value
5	557.17	412.26	1722.03	1721.64	11.07	5	3.55	4.29	3.80	1.35	11.07
10	975.85	762.70	2671.13	2906.49	18.31	10	8.26	7.53	13.16	10.17	18.31
15	1192.88	1087.96	3434.51	3849.18	25.00	15	9.97	9.12	17.60	19.12	25.00
20	1430.56	1404.29	4117.37	4500.64	31.41	20	11.27	12.38	24.02	29.27	31.41
50	3167.19	3137.33	7234.72	7743.76	67.50	50	45.86	44.97	90.65	51.34	67.50
100	5164.17	5240.23	11234.36	13131.06	124.34	100	102.40	70.91	178.10	95.96	124.34
150	6445.61	6632.65	13539.18	17527.54	179.58	150	168.24	106.01	246.26	162.45	179.58
200	7350.44	7601.17	15641.54	21028.11	233.99	200	234.11	131.46	316.12	228.84	233.99

The table reports the Ljung-Box statistic on trade durations (left side) and standardized residuals of trade duration (right side) for different stocks at different lags. The column of Lags is the number of lags we use to compute the statistic and the C\_Value is the critical value for the corresponding lags.

Table 4: Estimated Results For Quote Durations

		MEO			MRK			RWE			TKA		
		Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat
Panel A	ln(dur(lag1))	0.299	0.004	<b>70.792</b>	0.309	0.006	<b>52.986</b>	0.301	0.004	<b>75.553</b>	0.199	0.005	<b>39.587</b>
	ln(dur(lag2))	0.041	0.004	<b>9.511</b>	0.131	0.004	<b>29.611</b>	0.127	0.004	<b>35.714</b>	0.023	0.004	<b>5.949</b>
	ln(dur(lag3))	0.165	0.003	<b>48.008</b>							0.017	0.004	<b>4.831</b>
	x/phi (lag1)	0.072	0.001	<b>57.179</b>	0.042	0.001	<b>31.149</b>	0.063	0.001	<b>51.643</b>	0.047	0.001	<b>42.350</b>
	x/phi (lag2)										0.021	0.001	<b>22.672</b>
	x/phi (lag3)										0.017	0.001	<b>17.146</b>
	x/phi (lag4)										0.016	0.001	<b>23.487</b>
	cons	-0.575	0.016	<b>-36.725</b>	-0.394	0.012	<b>-31.852</b>	-1.110	0.013	<b>-83.628</b>	-1.821	0.017	<b>-106.145</b>
Panel B	x/phi (lag1)*dummy	0.070	0.004	<b>19.370</b>	0.069	0.005	<b>13.471</b>	0.075	0.004	<b>18.651</b>	0.071	0.003	<b>25.239</b>
	trade_dur/expected_trade_dur	0.154	0.002	<b>70.405</b>	0.269	0.003	<b>107.065</b>	0.267	0.002	<b>108.278</b>	0.254	0.002	<b>105.179</b>
	lag_trade_dur/expected_trade_dur	0.037	0.003	<b>12.023</b>	-0.001	0.002	<b>-0.584</b>	0.026	0.003	<b>8.439</b>	0.091	0.003	<b>34.055</b>
	lag_expected_trade_dur	0.108	0.005	<b>23.294</b>	-0.064	0.003	<b>-18.381</b>	0.202	0.004	<b>46.566</b>	0.226	0.004	<b>51.182</b>
	delta spread	-10.671	0.176	<b>-60.510</b>	-12.345	0.187	<b>-66.007</b>	-13.301	0.116	<b>-114.931</b>	-8.662	0.090	<b>-95.923</b>
	lev.spread	-4.573	0.098	<b>-46.586</b>	-1.560	0.109	<b>-14.254</b>	-9.825	0.116	<b>-84.423</b>	-5.026	0.106	<b>-47.236</b>
	square root(vol)	-0.072	3.72E-04	<b>-193.950</b>	-0.065	0.001	<b>-125.819</b>	-0.040	1.82E-04	<b>-219.646</b>	-0.029	1.69E-04	<b>-170.569</b>
	abs(sign.vol)	9.73E-05	1.69E-06	<b>57.550</b>	-3.51E-05	2.80E-06	<b>-12.525</b>	2.82E-05	1.10E-06	<b>25.673</b>	2.96E-05	7.10E-07	<b>41.690</b>
	Back_quote_dur	6.68E-04	1.94E-04	<b>3.444</b>	0.004	1.14E-04	<b>35.804</b>	0.010	2.83E-04	<b>35.118</b>	0.001	3.08E-04	<b>2.637</b>
	$\Delta$ XLm	-0.074	0.002	<b>-36.090</b>	-0.033	0.002	<b>-17.847</b>	-0.023	0.002	<b>-14.136</b>	0.012	0.003	<b>4.697</b>
Panel C	Dummy1	0.193	0.011	<b>17.960</b>	-0.390	0.012	<b>-33.898</b>	0.531	0.010	<b>55.626</b>	0.173	0.010	<b>17.074</b>
	Dummy2	0.139	0.007	<b>19.935</b>	-0.577	0.009	<b>-65.518</b>	-0.018	0.006	<b>-3.088</b>	-0.258	0.007	<b>-36.187</b>
	Dummy3	-0.067	0.006	<b>-11.290</b>	-0.115	0.008	<b>-15.127</b>	-0.055	0.006	<b>-8.858</b>	-0.231	0.007	<b>-31.943</b>
	Dummy4	-0.128	0.006	<b>-20.111</b>	-0.141	0.009	<b>-15.891</b>	-0.003	0.007	<b>-0.506</b>	-0.266	0.010	<b>-27.010</b>
	Dummy5	-0.097	0.006	<b>-15.359</b>	-0.131	0.009	<b>-15.115</b>	0.021	0.007	<b>2.790</b>	-0.020	0.009	<b>-2.186</b>
	Dummy6	-0.298	0.007	<b>-43.491</b>	-0.329	0.007	<b>-45.212</b>	-0.047	0.008	<b>-6.227</b>	-0.201	0.011	<b>-17.957</b>
	Dummy7	0.049	0.005	<b>9.788</b>	0.101	0.008	<b>12.343</b>	-0.089	0.006	<b>-13.938</b>	0.081	0.008	<b>9.593</b>
	Dummy8	-0.203	0.005	<b>-41.558</b>	0.015	0.008	<b>1.795</b>	-0.245	0.006	<b>-44.364</b>	-0.268	0.007	<b>-36.236</b>

The table reports the estimated results for quote durations. Panel A presents the results for Log-ACD part, Panel B includes the results for the exogenous variables and Panel C consists of the results for dummy variables. Bold entries indicate the coefficients are significant at a 5% level.

Table 5: Ljung-Box Statistics For Quote Durations and Deduced Standardized Residuals

Ljung-Box Statistics on Trade Durations						Ljung-Box Statistics on Corresponding Standardized Residuals					
Lags	MEO	MRK	RWE	TKA	C_Value	Lags	MEO	MRK	RWE	TKA	C_Value
5	120.09	34.11	297.78	492.92	11.07	5	10.43	5.92	8.1	8.12	11.070
10	134.80	42.80	369.75	611.70	18.31	10	15.54	15.96	9.87	49.33	18.307
15	164.76	52.64	392.25	687.89	25.00	15	24.77	23.1	11.25	72.68	24.996
20	181.16	63.90	424.92	727.08	31.41	20	45.75	32.93	12.38	116.32	31.410
50	278.87	117.19	650.20	952.42	67.50	50	74.21	47.55	53.54	251.1	67.505
100	408.87	169.94	1019.39	1373.74	124.34	100	151.98	91.85	131.14	345.85	124.342
150	559.91	212.97	1200.21	1623.94	179.58	150	220.29	128.82	178.94	441.45	179.581
200	664.12	257.83	1373.35	2121.54	233.99	200	278.68	172.9	217.33	518.5	233.994

The table reports the Ljung-Box statistic on quote durations (left side) and standardized residuals of quote duration (right side) for different stocks at different lags. The column of Lags is the number of lags we use to compute the statistic and the C\_Value is the critical value for the corresponding lags.

Table 6: Estimated Results For Liquidity Activity Factor of  $XLM^{25000}$

		MEO			MRK			RWE			TKA		
		Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat
Panel A	$g(t-1)$	0.661	0.084	<b>7.837</b>	0.848	0.053	<b>15.933</b>	0.762	0.041	<b>18.459</b>	0.954	0.016	<b>59.194</b>
	$activity(t-1)$	0.807	0.042	<b>19.351</b>	0.892	0.049	<b>18.153</b>	0.776	0.034	<b>22.695</b>	0.399	0.029	<b>13.564</b>
	$activity(t-2)$	-0.359	0.102	<b>-3.534</b>	-0.657	0.087	<b>-7.514</b>	-0.410	0.062	<b>-6.651</b>	-0.345	0.033	<b>-10.517</b>
	cons	-2.023	0.107	<b>-18.962</b>	-2.031	0.143	<b>-14.196</b>	-1.479	0.102	<b>-14.492</b>	-1.384	0.101	<b>-13.745</b>
Panel B	Expect Trade Dur	0.269	0.035	<b>7.735</b>	0.278	0.033	<b>8.529</b>	0.220	0.029	<b>7.471</b>	0.143	0.025	<b>5.852</b>
	Expect Quote Dur	-0.097	0.020	<b>-4.934</b>	-0.086	0.021	<b>-4.117</b>	-0.124	0.022	<b>-5.611</b>	-0.016	0.013	-1.245
	delta spread	3.633	0.649	<b>5.601</b>	6.324	1.003	<b>6.305</b>	6.341	0.859	<b>7.378</b>	1.433	0.565	<b>2.535</b>
	lev.spread	2.375	1.038	<b>2.287</b>	2.831	1.477	1.918	6.284	1.327	<b>4.737</b>	6.126	0.854	<b>7.176</b>
	square root(vol)	0.026	0.003	<b>9.192</b>	0.035	0.004	<b>8.876</b>	0.004	0.002	<b>2.132</b>	0.008	0.001	<b>7.517</b>
	abs(sign.vol)	1.52E-05	1.82E-05	0.834	4.05E-05	3.13E-05	1.293	-7.21E-06	0.000	-0.755	-1.54E-05	4.34E-06	<b>-3.541</b>
	Back_quote_dur	0.006	7.36E-04	<b>8.027</b>	0.004	6.40E-04	<b>7.010</b>	0.006	1.19E-03	<b>4.959</b>	0.005	9.99E-04	<b>4.526</b>
$\Delta XLM$	0.037	0.013	<b>2.945</b>	0.062	0.012	<b>5.162</b>	0.071	0.009	<b>7.852</b>	0.076	0.015	<b>5.129</b>	
Panel C	Dummy1	0.190	0.134	1.418	0.028	0.134	0.211	-0.051	0.089	-0.574	0.174	0.071	<b>2.464</b>
	Dummy2	-0.084	0.087	-0.966	-0.205	0.090	<b>-2.272</b>	-0.097	0.065	-1.500	0.038	0.052	0.729
	Dummy3	-0.096	0.084	-1.152	-0.042	0.093	-0.452	-0.075	0.066	-1.141	0.104	0.056	1.862
	Dummy4	-0.193	0.093	<b>-2.067</b>	-0.224	0.097	<b>-2.310</b>	-0.140	0.066	<b>-2.132</b>	-0.065	0.064	-1.013
	Dummy5	-0.133	0.089	-1.503	-0.192	0.107	-1.799	-0.243	0.082	<b>-2.981</b>	-0.020	0.073	-0.273
	Dummy6	-0.096	0.093	-1.036	-0.103	0.098	-1.053	-0.219	0.077	<b>-2.844</b>	0.011	0.074	0.153
	Dummy7	-0.122	0.074	-1.659	-0.315	0.095	<b>-3.322</b>	-0.136	0.067	<b>-2.026</b>	-0.083	0.064	-1.298
	Dummy8	-0.034	0.064	-0.534	-0.112	0.088	-1.270	-0.033	0.055	-0.605	0.084	0.051	1.650

The table reports the estimated results for Activity factor. Panel A presents the results for GLARMA part, Panel B includes the results for the exogenous variables and Panel C consists of the results for dummy variables. Bold entries indicate the coefficients are significant at a 5% level.

Table 7: Ljung-Box Statistics For Activity Factor of  $XLM^{25000}$  and Deduced Standardized Residuals

Lags	Ljung-Box Statistics on Trade Durations					Ljung-Box Statistics on Corresponding Standardized Residuals					
	MEO	MRK	RWE	TKA	C_Value	Lags	MEO	MRK	RWE	TKA	C_Value
5	887.04	754.79	1127.06	358.71	11.07	5	3.08	4.51	3.68	8.32	11.07
10	979.11	850.97	1272.35	441.53	18.31	10	5.36	6.64	6.72	12.8	18.31
15	1007.67	923.88	1322.27	512.89	25.00	15	8.12	16.48	11.28	15.51	25.00
20	1022.06	957.06	1329.06	540.47	31.41	20	11.45	23.48	16.43	21.91	31.41
50	1233.96	1056.15	1417.73	662.32	67.50	50	54.39	61.57	50.13	54.5	67.50
100	1453.23	1163.59	1502.42	813.11	124.34	100	99.63	116.12	108.88	111.38	124.34
150	1555.10	1274.25	1538.54	917.70	179.58	150	147.95	177.37	145.24	151.97	179.58
200	1665.43	1332.34	1586.24	994.88	233.99	200	203.86	234.75	190.94	189.69	233.99

The table reports the Ljung-Box statistic on activity factor (left side) and standardized residuals of activity factor (right side) for different stocks at different lags. The column of Lags is the number of lags we use to compute the statistic and the C\_Value is the critical value for the corresponding lags.

Table 8: Estimated Results For Direction Factor of  $XLM^{25\ 000}$

		MEO			MRK			RWE			TKA		
		Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat
Panel A	$g(t-1)$	1.06	0.07	<b>14.77</b>	0.90	0.10	<b>9.28</b>	0.94	0.07	<b>13.97</b>	0.50	0.02	<b>26.64</b>
	$g(t-2)$	-0.30	0.06	<b>-4.72</b>	-0.14	0.08	-1.65	-0.30	0.09	<b>-3.45</b>	0.03	0.02	1.61
	$g(t-3)$							0.08	0.04	1.73			
	$direction(t-1)$	-0.57	0.05	<b>-10.77</b>	-0.55	0.06	<b>-9.11</b>	-0.67	0.04	<b>-14.95</b>	-1.70	0.04	<b>-42.55</b>
	cons	0.04	0.11	0.38	-0.13	0.13	-0.98	-0.06	0.09	-0.71	0.18	0.09	<b>2.09</b>
Panel B	Expect Quote	-0.01	0.01	-1.25	0.02	0.01	1.51	0.01	0.01	1.32	0.02	0.01	<b>2.52</b>
	Quote Dur												
	delta spread	4.84	1.03	<b>4.70</b>	3.05	1.12	<b>2.73</b>	2.40	0.95	<b>2.54</b>	2.75	0.73	<b>3.79</b>
	lev.spread	-5.37	1.37	<b>-3.91</b>	-1.14	1.91	-0.60	-3.00	1.64	-1.83	-7.11	0.98	<b>-7.24</b>
	square root(vol)	1.15E-02	4.07E-03	<b>2.82</b>	7.88E-04	0.01	0.15	4.18E-03	2.41E-03	1.73	7.14E-03	1.82E-03	<b>3.92</b>
	abs(sign.vol)	-3.51E-05	2.52E-05	-1.39	6.55E-05	4.01E-05	1.64	-4.53E-06	1.34E-05	-0.34	-1.63E-05	7.23E-06	<b>-2.25</b>
Back_quote_dur	-2.97E-03	8.40E-04	<b>-3.54</b>	-1.15E-03	6.75E-04	-1.70	-0.01	1.25E-03	-4.61	-3.49E-03	1.62E-03	<b>-2.15</b>	
	$\Delta XLM$	-0.33	0.02	<b>-14.10</b>	-0.22	0.02	<b>-9.52</b>	-0.29	0.02	<b>-14.65</b>	-0.13	0.02	<b>-6.82</b>
Panel C	Dummy1	0.11	0.17	0.62	-0.12	0.17	-0.71	0.23	0.12	1.88	0.06	0.11	0.61
	Dummy2	0.01	0.13	0.05	0.20	0.12	1.64	0.12	0.09	1.33	-0.10	0.09	-1.10
	Dummy3	0.07	0.12	0.60	0.07	0.12	0.56	0.22	0.09	<b>2.47</b>	0.06	0.09	0.61
	Dummy4	0.16	0.13	1.18	0.03	0.13	0.20	0.14	0.09	1.54	-0.13	0.11	-1.18
	Dummy5	0.05	0.13	0.36	0.10	0.15	0.69	0.10	0.11	0.94	0.00	0.12	0.02
	Dummy6	0.02	0.13	0.17	-0.08	0.13	-0.63	0.20	0.10	1.91	-0.16	0.12	-1.32
	Dummy7	0.10	0.11	0.92	0.09	0.13	0.71	0.15	0.09	1.65	-0.22	0.11	-1.93
	Dummy8	0.09	0.10	0.90	-0.04	0.12	-0.34	0.09	0.08	1.17	-0.08	0.09	-0.83

The table reports the estimated results for Direction factor. Panel A presents the results for GLARMA part, Panel B includes the results for the exogenous variables and Panel C consists of the results for dummy variables. Bold entries indicate the coefficients are significant at a 5% level.

Table 9: Ljung-Box Statistics For Direction Factor of  $XLM^{25\ 000}$  and Deduced Standardized Residuals

Ljung-Box Statistics on Trade Durations						Ljung-Box Statistics on Corresponding Standardized Residuals					
Lags	MEO	MRK	RWE	TKA	C_Value	Lags	MEO	MRK	RWE	TKA	C_Value
5	869.93	720.97	1963.59	3051.37	11.07	5	6.14	2.79	11.05	3.59	11.07
10	871.34	722.21	1977.68	3059.01	18.31	10	11.95	4.73	20.31	10.78	18.31
15	888.31	729.77	1987.41	3060.47	25.00	15	20.7	12.29	22.22	18.23	25.00
20	897.14	733.64	1996.95	3066.62	31.41	20	24.56	17.98	29.49	22.98	31.41
50	950.96	780.98	2020.14	3114.22	67.50	50	50.02	51.02	53.49	60.31	67.50
100	1004.69	871.68	2073.82	3238.26	124.34	100	84.32	109.29	90.04	143.35	124.34
150	1051.74	966.96	2148.45	3272.82	179.58	150	138.29	169.6	139.51	182.42	179.58
200	1114.54	1045.62	2194.39	3392.66	233.99	200	205.41	223.39	193.7	232.18	233.99

The table reports the Ljung-Box statistic on direction factor (left side) and standardized residuals of direction factor (right side) for different stocks at different lags. The column of Lags is the number of lags we use to compute the statistic and the C\_Value is the critical value for the corresponding lags.

Table 10: Estimated Results For Liquidity Size Factor of  $XLM^{25\ 000}$

		MEO			MRK			RWE			TKA		
		Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat
Panel A	g(t-1)	0.70	0.16	<b>4.36</b>	0.79	0.08	<b>9.60</b>	0.59	0.15	<b>3.95</b>	-0.01	0.36	-0.02
	g(t-2)	-0.16	0.02	<b>-7.45</b>	-0.13	0.02	<b>-7.90</b>	0.14	0.09	1.61	0.18	0.07	<b>2.48</b>
	Size (t-1)	0.07	0.04	1.77	0.07	0.03	<b>2.66</b>	-0.22	0.01	<b>-16.83</b>	-0.49	0.04	<b>-13.92</b>
	Size (t-2)							0.11	0.03	<b>3.37</b>	0.00	0.18	-0.01
	cons	0.39	0.10	<b>4.04</b>	0.28	0.11	<b>2.44</b>	0.42	0.08	<b>5.43</b>	1.90	0.08	<b>23.29</b>
Panel B	Expect QuoteQuote Dur	0.00	0.00	0.56	0.00	0.00	-1.87	0.00	0.00	-0.97	0.00	0.00	-1.50
	Direction(t)	-0.07	0.03	<b>-2.51</b>	-0.06	0.03	<b>-2.31</b>	-0.10	0.02	<b>-5.27</b>	-0.15	0.04	<b>-4.02</b>
	Direction(t-1)	-0.01	0.05	-0.12	0.06	0.04	1.26	0.09	0.03	<b>2.71</b>	0.15	0.06	<b>2.58</b>
	delta spread	1.03	0.53	1.93	1.16	0.87	1.33	1.97	0.71	<b>2.76</b>	0.25	0.37	0.68
	lev.spread	0.96	1.05	0.92	0.22	1.36	0.16	0.39	1.24	0.31	-1.76	0.75	<b>-2.36</b>
	square root(vol)	0.00	0.00	-0.67	0.00	0.00	-0.71	0.00	0.00	1.31	-0.01	0.00	<b>-3.16</b>
	abs(sign.vol)	0.00	0.00	-0.08	0.00	0.00	-0.19	0.00	0.00	-0.74	0.00	0.00	-1.41
	Back_quote_dur	0.00	0.00	1.58	0.00	0.00	1.60	0.01	0.00	<b>5.35</b>	0.01	0.00	<b>4.38</b>
$\Delta XLM$	-0.01	0.02	-0.28	-0.03	0.02	-1.95	-0.07	0.01	<b>-5.28</b>	-0.15	0.03	<b>-4.30</b>	
Panel C	Dummy1	-0.13	0.13	-1.02	-0.17	0.13	-1.33	-0.16	0.09	-1.83	-0.54	0.10	<b>-5.26</b>
	Dummy2	-0.07	0.10	-0.72	-0.10	0.09	-1.10	-0.05	0.06	-0.71	-0.26	0.10	<b>-2.72</b>
	Dummy3	-0.12	0.10	-1.21	-0.04	0.09	-0.43	-0.05	0.06	-0.76	-0.19	0.10	-1.94
	Dummy4	-0.16	0.10	-1.59	-0.03	0.10	-0.32	-0.09	0.06	-1.41	-0.29	0.12	<b>-2.51</b>
	Dummy5	0.06	0.10	0.57	-0.10	0.11	-0.87	-0.02	0.08	-0.21	-0.31	0.12	<b>-2.51</b>
	Dummy6	-0.10	0.10	-0.95	-0.09	0.10	-0.99	0.03	0.08	0.36	-0.20	0.13	-1.50
	Dummy7	-0.03	0.09	-0.33	-0.02	0.10	-0.17	-0.04	0.07	-0.58	-0.12	0.12	-1.03
	Dummy8	-0.06	0.08	-0.72	-0.04	0.09	-0.43	-0.01	0.06	-0.23	-0.14	0.10	-1.41

The table reports the estimated results for Size factor. Panel A presents the results for GLARMA part, Panel B includes the results for the exogenous variables and Panel C consists of the results for dummy variables. Bold entries indicate the coefficients are significant at a 5% level.

Table 11: Ljung-Box Statistics For Size Factor of  $XLM^{25\ 000}$  and Deduced Standardized Residuals

Ljung-Box Statistics on Trade Durations						Ljung-Box Statistics on Corresponding Standardized Residuals					
Lags	MEO	MRK	RWE	TKA	C_Value	Lags	MEO	MRK	RWE	TKA	C_Value
5	247.48	297.58	1406.47	889.69	11.07	5	2.73	4.59	18.64	21.23	11.07
10	252.01	330.53	1592.66	1054.64	18.31	10	2.96	11.11	23.22	27.89	18.31
15	263.01	343.90	1624.90	1107.16	25.00	15	12.67	13.86	24.03	36.5	25.00
20	264.68	350.70	1655.55	1124.97	31.41	20	14.05	19.96	32.73	39.11	31.41
50	301.70	383.77	1701.23	1218.48	67.50	50	49.69	45.29	75.52	68.95	67.50
100	335.13	429.84	1774.36	1266.30	124.34	100	80.15	88.45	116.21	114.54	124.34
150	389.70	497.10	1814.41	1301.94	179.58	150	141.83	132.88	163.7	155.42	179.58
200	458.14	550.77	1909.40	1358.22	233.99	200	205.49	177.29	219.94	199.67	233.99

The table reports the Ljung-Box statistic on size factor (left side) and standardized residuals of size factor (right side) for different stocks at different lags. The column of Lags is the number of lags we use to compute the statistic and the C\_Value is the critical value for the corresponding lags.

Table 12: Effect of Key Exogenous Variables on XLM

Panel A: Effects on $XLM^{5000}$ and $XLM^{25000}$ changes				
	Activity	Direction	Size (magnitude)	Total Effect on Ex-ante Liquidity
Delta spread	+ (4,4)	+ (4,4)	+ (4,1)	If the previous spread increases (less liquid), the liquidity is likely to decrease, the size of this decrease in liquidity is likely to be small.
Lag XLM	+ (4,4)	- (3,4)	- (2,2)	If the market is less liquid, the following liquidity is likely to increase and the size of this increase will be large.
Direction(t)			- (3,4)	Leverage effect: the size of liquidity decrease is larger than liquidity increase.
Panel B: Effect on $XLM^{25000}$ changes				
	Activity	Direction	Size	Total Effect on Ex-ante Liquidity
Ave_spread	+ (3)	- (2)	- (1)	If Ave_Spread increases, the liquidity is likely to increase, the size of this increase is likely to be large.
Square root(vol)	+ (4)	+ (2)	- (1)	If the trade volume increases, the liquidity is likely to decrease, the size of this decrease is likely to be large.
Abs(sign.vol)	- (1)	- (1)	#	If imbalance of trade increases, the liquidity is likely to remain the same, if the liquidity changes, it will increase and imbalance does not affect the size of the change.
Panel C: Effect on $XLM^{5000}$ changes				
	Activity	Direction	Size	Total Effect on Ex-ante Liquidity
Ave_spread	+ (3)	- (2)	+ (2)	If Ave_Spread increases, the liquidity is likely to increase, the size of this increase is likely to be moderate.
Square root(vol)	- (3)	#	+ (4)	If the trade volume increases, the liquidity is likely to remain the same, if the liquidity changes, the size of this change is likely to be small.
Abs(sign.vol)	- (1)	#	#	If imbalance of trade increases, the liquidity is likely to remain the same, the imbalance of trades does not affect the size of the change.

The table summarizes the effect of various exogenous variables on the  $XLM^{5000}$  and  $XLM^{25000}$  changes, # relates to a non-significant estimated coefficient at a 5% level. In panel A, + (4,4) means that four stocks have significantly positive coefficients for  $XLM^{5000}$  changes and four for  $XLM^{25000}$  changes.



## References

- Aitken, M., Comerton-Forde, C., 2003. How should liquidity be measured? *Pacific-Basin Finance Journal* 11 (1), 45–59.
- Bauwens, L., Giot, P., 2000. The Logarithmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks. *Annals of Economics and Statistics / Annales d'Économie et de Statistique* (60), 117–149.
- Beltran-Lopez, H., Giot, P., Grammig, J., 2009. Commonalities in the order book. *Financial Markets and Portfolio Management* 23 (3), 209–242.
- Beltran-Lopez, H., Grammig, J., Menkveld, A. J., 2011. Limit order books and trade informativeness. *The European Journal of Finance*, 1–23.
- Biais, B., Wooley, P., 2011. High frequency trading. Unpublished working paper. University of Toulouse, Industrial Economics Institute, Toulouse, France.
- Bilodeau, Y., 2013. Xetraparser [computer software]. HEC Montréal.
- Black, F., 1971. Toward a fully automated stock exchange. *Financial Analysts Journal* 27 (4), 28–44.
- Bloomfield, R., O'hara, M., Saar, G., 2015. Hidden liquidity: Some new light on dark trading. *The Journal of Finance* 70 (5), 2227–2274.
- Boehmer, E., Saar, G., Yu, L., 2005. Lifting the veil: An analysis of pre-trade transparency at the nyse. *The Journal of Finance* 60 (2), 783–815.
- Brogaard, J., Hendershott, T., Riordan, R., 2014. High-frequency trading and price discovery. *Review of Financial Studies*.
- Coppejans, M., Domowitz, I., Madhavan, A., 2004. Resiliency in an automated auction. Working paper, ITG Group.
- Coughenour, J., Shastri, K., 1999. Symposium on market microstructure: A review of empirical research. *Financial Review* 34 (4), 1–27.
- Cox, D. R., Gudmundsson, G., Lindgren, G., Bondesson, L., Harsaae, E., Laake, P., Juselius, K., Lauritzen, S. L., Jan. 1981. *Statistical Analysis of Time Series: Some Recent Developments [with Discussion and Reply]*. *Scandinavian Journal of Statistics* 8 (2), 93–115.
- Demsetz, H., 1968. The Cost of Transacting. *Quarterly Journal of Economics* 82 (1), 33–53.
- Dionne, G., Duchesne, P., Pacurar, M., 2009. Intraday value at risk (ivar) using tick-by-tick data with application to the toronto stock exchange. *Journal of Empirical Finance* 16 (5), 777–792.
- Dionne, G., Pacurar, M., Zhou, X., 2015. Liquidity-adjusted intraday value at risk modeling and risk management: An application to data from deutsche borse. *Journal of Banking and Finance* 59, 202 – 219.
- Domowitz, I., Hansch, O., Wang, X., 2005. Liquidity commonality and return co-movement. *Journal of Financial Markets* 8 (4), 351–376.
- Easley, D., O'Hara, M., 1987. Price, trade size, and information in securities markets. *Journal of Financial Economics* 19 (1), 69–90.

- Engle, R. F., Lange, J., 2001. Predicting VNET: A model of the dynamics of market depth. *Journal of Financial Markets* 4 (2), 113–142.
- Engle, R. F., Lunde, A., 2003. Trades and Quotes: A Bivariate Point Process. *Journal of Financial Econometrics* 1 (2), 159–188.
- Engle, R. F., Russell, J. R., 1998. Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica* 66 (5), 1127–1162.
- Foucault, T., 1999. Order flow composition and trading costs in a dynamic limit order market. *Journal of Financial Markets* 2 (2), 99–134.
- Foucault, T., Kadan, O., Kandel, E., 2005. Limit Order Book as a Market for Liquidity. *Review of Financial Studies* 18 (4), 1171–1217.
- Giot, P., Grammig, J., 2006. How large is liquidity risk in an automated auction market? *Empirical Economics* 30 (4), 867–887.
- Glosten, L. R., Harris, L. E., 1988. Estimating the Components of the Bid Ask Spread. *Journal of Financial Economics* 21 (1), 123.
- Goodhart, C. A. E., O’Hara, M., 1997. High frequency data in financial markets: Issues and applications. *Journal of Empirical Finance* 4 (2-3), 73–114.
- Hasbrouck, J., 1991. Measuring the Information Content of Stock Trades. *The Journal of Finance* 46 (1), 179–207.
- Hasbrouck, J., 1996. Order characteristics and stock price evolution An application to program trading. *Journal of Financial Economics* 41 (1), 129–149.
- Hasbrouck, J., Saar, G., 2013. Low-latency trading. *Journal of Financial Markets* 16 (4), 646 – 679, high-Frequency Trading.
- Hausman, J. A., Lo, A. W., MacKinlay, A. C., 1992. An ordered probit analysis of transaction stock prices. *Journal of Financial Economics* 31 (3), 319–379.
- Hendershott, T., Jones, C., Menkveld, A., 2011. Does algorithmic trading improve liquidity? *The Journal of Finance* 66 (1), 1–33.
- Hendershott, T., Riordan, R., 2013. Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis* 48, 1001–1024.
- Irvine, P. J., Benston, G. J., Kandel, E., 2000. Liquidity Beyond the Inside Spread: Measuring and Using Information in the Limit Order Book. SSRN eLibrary.
- Jones, C., 2012. What do we know about high frequency trading? Unpublished working paper. Columbia University.
- Lee, C. M. C., Ready, M. J., Jun. 1991. Inferring Trade Direction from Intraday Data. *The Journal of Finance* 46 (2), 733–746.
- Madhavan, A., 2000. Market microstructure: A survey. *Journal of Financial Markets* 3 (3), 205–258.
- Manganelli, S., 2005. Duration, volume and volatility impact of trades. *Journal of Financial Markets* 8 (4), 377–399.

- McCulloch, R. E., Tsay, R. S., 2001. Nonlinearity in High-Frequency Financial Data and Hierarchical Models. *Studies in Nonlinear Dynamics & Econometrics* 5 (1), 1–17.
- Menkveld, A. J., 2013. High frequency trading and the new market makers. *Journal of Financial Markets* 16 (4), 712 – 740, high-Frequency Trading.
- O’Hara, M., 2015. High frequency market microstructure. *Journal of Financial Economics* 116 (2), 257 – 270.
- Parlour A., C., 1998. Price Dynamics in Limit Order Markets. *The Review of Financial Studies* 11 (4), 789.
- Riordan, R., Storkenmaier, A., 2012. Latency, liquidity and price discovery. *Journal of Financial Markets* 15 (4), 416 – 437.
- Rogers, L. C. G., Zane, O., 1998. Designing and estimating models of high-frequency data. Unpublished paper.
- Russell, J., Engle, R., 2005. A Discrete-State Continuous-Time Model of Financial Transactions Prices and Times: The Autoregressive Conditional Multinomial-Autoregressive Conditional Duration Model. *Journal of Business & Economic Statistics* 23 (2), 166.
- Rydberg, T. H., Shephard, N., Mar. 2003. Dynamics of Trade-by-Trade Price Movements: Decomposition and Models . *Journal of Financial Econometrics* 1 (1), 2–25.
- Shephard, N., 1995. Generalized linear autoregressions. Working paper, Nuffield College, University of Oxford.
- Viswanathan, S., Wang, J. J. D., 2002. Market architecture: limit-order books versus dealership markets. *Journal of Financial Markets* 5 (2), 127–167.

## Appendix

### 1. Trade Duration Factor

Panel A of Table 2 presents the estimated results of the lagged dependent variables for all the stocks in the sample. The ACD-related models are based on the ARMA structure, meaning that the determinant parts of the dynamics of durations are specified by the lagged duration and lagged error term <sup>1</sup>. As we expected, the trade durations of all stocks are highly autocorrelated. Taking the models' efficiency and parsimony into consideration, the number of lagged trade durations varies from one to three. The sum of the coefficients of lagged trade durations for all stocks is around 0.9, which confirms a very high degree of persistence in autocorrelation. TKA has the most persistent trade durations among all stocks and has been specified by Log-ACD(3,2) dynamics. Regarding the error terms for all stocks, the parameters *Gamma1* (ranges from 0.512 to 0.974) and *Gamma2* (ranges from 0.350 to 0.780) are small, indicating that the distribution of the random part is abrupt and near to zero; that is, expected trade duration shocks are very small.

Considering the exogenous variables mentioned above, we put six such variables in the trade duration equations: *DeltaSpread*, *AveSpread*, *SquareRoot(vol)*, *Abs(sign.vol)*, *BackQuote duration*, and  $\Delta XLM$ . Panel B of Table 2 reports the estimation results. Consistent with Engle and Lunde (2003), for all the stocks, a positive *DeltaSpread*, increases the trade duration. The parameters are all positive and significant at a 5% level. Intuitively, when the stocks become less liquid, traders will slow their trading intensities, and trading durations will become longer. Contrary to Engle and Lunde (2003), we find a positive sign for *AveSpread* coefficient. That is, the higher the *AveSpread* is, the longer the trade durations are. The difference between *DeltaSpread* and *AveSpread* is that the former captures the short-run spread change and the latter relates to the long-term spread change. It reveals that trade intensity is sensitive for both short-run and long-run spread change. When the stocks become less liquid, traders slow down their activities and then wait for the moment when liquidity increases.

Another important exogenous variable is volume-related. For all stocks, the coefficients of *SquareRoot(vol)* are all negative and significant at a 5% level. This means that large trades will generate higher trading intensity and shorter trade durations. According to Easley and O'Hara (1987), large trades are likely to be related to information trade because the informed traders try to exploit their information advantage by increasing both the trading intensity and trading quantity. Another volume-related variable is *Abs(sign.vol)* which measures the volume imbalance (in shares) of the last ten trades (a higher value means a higher level of imbalance and a zero value means a completely balanced trading activity). All stocks have small positive

---

<sup>1</sup> Regarding the model selection, the first criteria is the Ljung-Box statistics, and the second one is the BIC criteria for the selection of number of lags.

coefficients, and only two of them are significant at a 5% level. The measure of imbalance also relates to information trading. Our results suggest that when the trades become imbalanced, the trade intensity will decrease. Presumably, when the non-informed traders observe an imbalanced trading history, they will then slow down their trading activity to protect themselves.

Regarding the *BackQuote duration*, we find that the effect of duration between the first quote and the following trade on the trade duration is not clear. Two stocks have a positive sign and two others have a negative sign. Moreover, three of them are significant at a 5% level. Intuitively, the trade activity should be positively correlated with quote activity. As a result, the coefficient should have a positive sign. Our results do not seem to be completely consistent with this intuition. One explanation might be that trade duration consists of quote duration while *BackQuote duration* depends on both trade duration and quote duration. Therefore, the explanatory power of *BackQuote duration* is also influenced by quote duration, which is censored. For instance, a longer *BackQuote duration* could be either a long trade duration or a shorter quote duration. The longer trade duration has a positive impact on trade duration, whereas a shorter quote duration has a negative one. Therefore, the sign of *BackQuote duration* on trade duration could be positive or negative.

The last exogenous variable is the change of liquidity measure XLM. Like *DeltaSpread*, a high XLM means a less liquid situation. If we follow the same argument as *DeltaSpread*, higher XLM will decrease the trade intensity and then generate a longer trade duration. Our estimated results show that  $\Delta XLM^{25000}$  of RWE and  $\Delta XLM^{5000}$  of MEO and TKA have positive estimated coefficients which are significant at a 5% level. This suggests that the  $\Delta XLM$  is likely to have a moderated effect on trade duration. However, if we combine the  $\Delta XLM$  result with that of *DeltaSpread*, it suggests that the trade activity is more sensitive to the first level of LOB and the quote activities beyond the first level are not likely to explain the trade duration dynamics.

The last part consists of dummy variables. Panel C of Table 2 reports the estimated results for these dummy variables. To make the results more visible, we piecewisely plot the coefficient of dummy variables in Figure 2. In line with previous literature, the trade durations do have a diurnal pattern. For all stocks, there is an inverse U-shaped pattern for trade durations.

## 2 Quote Duration Factor

Panel A of Table 4 reports the estimated results of lagged dependent variables for all stocks. Similar to trade durations, quote durations are also highly persistent. The number of lagged values varies from two to three and the number of lagged error terms varies from one to four. TKA features the most persistent dynamics and takes a Log-ACD(3,4) model to fit its dynamics. Comparing the Log-ACD model to the traditional ARMA structure, the sum of the coefficient of the “AR” part in the quote duration equation varies from 0.239 to 0.505. The relatively small coefficients can be explained by the fact that the start point of our quote duration is the trade timestamp. When there is more than one update between two trades, only the first one is used to calculate the quote duration. To some extent, this sampling “deletes” some autocorrelation

in quote durations. Regarding the coefficients of the error term, the estimated coefficients vary from 0.042 to 0.072, confirming the sampling effect in autocorrelation.

The censored effect is captured by the product of the censored dummy variable and the lagged error term. The censored effect is positive and significant at a 5% level for all stocks. The implication is that if the last quote duration is censored, the next quote duration is likely to be longer and censored again. Trade duration related variables consist of current error term, lagged error term and lagged expected trade duration. The estimated results show that the current error term and expected trade duration have a positive and significant effect on the quote durations. It suggests that when the (current and lagged) trade duration innovations and expected trade durations are high, the quote duration will become longer. Intuitively, as quote activity adjusts to trade activity, trade duration and quote duration are correlated positively.

The spread-related variables, *DeltaSpread* and *AveSpread*, have a negative impact on the quote durations, which is opposite to the trade duration. The coefficients are negative and significant at 5% level for all stocks. Spread variables are very important in explaining the trade activity and quote activity. It suggests that when the spread is large, the market order traders and LOB traders react differently. Market order traders slow down their trading speed when observing an increasing spread, whereas LOB traders speed up to update their price or quantity of limit orders. For an intraday LOB trader, a major concern is adverse selection. By attentively monitoring the change of spread, LOB traders attempt to avoid this risk by updating their quote rapidly.

The effect of volume-related variables, *SquareRoot(vol)* and *Abs(sign.vol)*, on quote duration is similar to that of trade duration. More specifically, large trades predict shorter quote durations. As mentioned above, large trades relate to informed trades. Informed traders exploit short-lived information by increasing trade intensity and trade quantity. When trades become more intensive, so do quote revisions. As a result, the quote durations become shorter. Regarding imbalanced measures *Abs(sign.vol)*, their coefficients are all positive but only two of them are significant. Compared to *SquareRoot(vol)*, *Abs(sign.vol)* is a long-run variable, as it consists of information over the last ten trades. In a high-frequency trading framework, trades and information are short-lived. The estimated results suggest that the short-run measure is more predictive than the long-run measure.

The *BackQuote duration* has a positive and significant impact on the following quote duration. A longer *BackQuote duration* can be either a low trade intensity or a high quote intensity. That is, a long trade duration or a short quote duration. Moreover, a long trade duration implies a long quote duration. The estimated results suggest that the long trade duration effect dominates the short quote duration effect. Concerning the effect of liquidity measure  $XLM^{25000}$ , empirically, the effect is negative and significant at a 5% level for three stocks. One stock has a positive significant sign. We obtain similar results for  $XLM^{5000}$ . Theoretically, there are arguments to support either sign. A high XLM reveals a more risky situation, especially a high adverse selection risk. LOB traders speed up their quote revisions to reduce this kind of risk. This

is more pronounced for less liquid stocks. On the other side, the increase of XLM will slow down the trade activity and then the quote activity. Panel C of Table 4 reports the results of estimation for dummy variables. We also find a seasonality effect for quote durations. However, because the quote durations are censored and the trade durations enter into the model of quote durations, the seasonality pattern for quote durations is not like that of trade durations. Figure 3 illustrates the intraday seasonality pattern for different stocks.

Table A.1: Estimated Results For Trade Durations

		MEO			MRK			RWE			TKA		
		Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat
Panel A	ln(dur(lag1))	0.889	0.039	<b>22.653</b>	0.801	0.040	<b>19.912</b>	0.984	0.025	<b>38.606</b>	1.185	0.037	<b>32.402</b>
	ln(dur(lag2))	-0.048	0.038	-1.264	0.130	0.056	<b>2.328</b>	-0.138	0.037	<b>-3.672</b>	-0.296	0.061	<b>-4.876</b>
	ln(dur(lag3))	0.073	0.024	<b>2.979</b>	0.024	0.038	0.625	0.109	0.027	<b>3.982</b>	0.092	0.032	<b>2.877</b>
	x/phi (lag1)	0.078	0.008	<b>9.954</b>	0.059	0.006	<b>10.001</b>	0.117	0.007	<b>17.430</b>	0.132	0.009	<b>15.061</b>
	x/phi (lag2)										-0.058	0.009	<b>-6.655</b>
	gammel	0.510	0.021	<b>24.458</b>	0.514	0.024	<b>21.677</b>	0.972	0.026	<b>37.086</b>	0.797	0.019	<b>42.138</b>
	gamme2	0.786	0.052	<b>15.108</b>	0.713	0.054	<b>13.221</b>	0.351	0.014	<b>25.738</b>	0.450	0.016	<b>27.413</b>
	cons	0.249	0.034	<b>7.404</b>	0.109	0.027	<b>4.105</b>	0.052	0.016	<b>3.294</b>	0.008	0.007	1.117
Panel B	delta spread	14.854	0.274	<b>54.199</b>	26.400	0.926	<b>28.520</b>	18.312	0.601	<b>30.472</b>	14.138	0.550	<b>25.727</b>
	lev.spread	1.145	0.189	<b>6.062</b>	1.224	0.225	<b>5.438</b>	0.807	0.175	<b>4.622</b>	0.508	0.087	<b>5.807</b>
	square root(vol)	-0.014	0.001	<b>-9.615</b>	-0.008	0.002	<b>-5.045</b>	-0.008	0.001	<b>-11.997</b>	-0.003	0.000	<b>-9.923</b>
	abs(sign.vol)	4.52E-06	4.96E-06	0.911	3.27E-06	7.03E-06	0.465	4.52E-06	1.63E-06	<b>2.779</b>	1.25E-06	5.48E-07	<b>2.286</b>
	Back_quote_dur	5.14E-04	5.54E-04	0.928	8.09E-04	3.23E-04	<b>2.504</b>	-0.002	0.001	<b>-3.277</b>	-0.002	4.06E-04	<b>-4.200</b>
	$\Delta$ XLM	0.009	0.004	<b>2.322</b>	-0.001	0.003	-0.228	0.004	0.003	1.238	0.017	0.005	3.144
Panel C	Dummy1	0.046	0.024	1.893	-0.111	0.020	<b>-5.483</b>	-0.002	0.011	-0.205	-0.022	0.006	<b>-3.697</b>
	Dummy2	-0.002	0.016	-0.125	-0.065	0.013	<b>-5.196</b>	0.048	0.008	<b>5.896</b>	-0.002	0.004	-0.583
	Dummy3	0.031	0.015	<b>2.056</b>	-0.015	0.011	-1.304	0.039	0.009	<b>4.539</b>	0.008	0.004	<b>2.039</b>
	Dummy4	0.064	0.018	<b>3.600</b>	-0.002	0.012	-0.159	0.041	0.009	<b>4.492</b>	0.024	0.005	<b>4.769</b>
	Dummy5	0.054	0.016	<b>3.369</b>	-0.008	0.014	-0.594	0.081	0.010	<b>7.744</b>	0.039	0.006	<b>6.662</b>
	Dummy6	0.052	0.017	<b>2.996</b>	-0.021	0.012	-1.729	0.068	0.009	<b>7.215</b>	0.032	0.007	<b>4.728</b>
	Dummy7	0.012	0.013	0.936	-0.002	0.011	-0.201	0.055	0.009	<b>6.500</b>	0.030	0.005	<b>5.381</b>
	Dummy8	-0.039	0.011	<b>-3.433</b>	-0.004	0.011	-0.373	-0.001	0.007	-0.144	0.005	0.004	1.226

The table reports the estimated results for trade durations. Panel A presents the results for Log-ACD part, Panel B includes the results for the exogenous variables and Panel C consists of the results for dummy variables. Bold entries indicate the coefficients are significant at a 5% level.

Table A.2: Ljung-Box Statistics For Trade Durations and Deduced Standardized Residuals

Lags	Ljung-Box Statistics on Trade Durations					Ljung-Box Statistics on Corresponding Standardized Residuals					
	MEO	MRK	RWE	TKA	C_Value	Lags	MEO	MRK	RWE	TKA	C_Value
5	557.17	412.26	1722.03	1721.64	11.07	5	2.61	4.28	4.21	1.21	11.07
10	975.85	762.70	2671.13	2906.49	18.31	10	7.32	7.58	13.98	10.77	18.31
15	1192.88	1087.96	3434.51	3849.18	25.00	15	8.78	9.17	18.38	19.08	25.00
20	1430.56	1404.29	4117.37	4500.64	31.41	20	10.23	12.36	25.23	28.91	31.41
50	3167.19	3137.33	7234.72	7743.76	67.50	50	45.75	44.23	92.44	51.75	67.50
100	5164.17	5240.23	11234.36	13131.06	124.34	100	101.21	70.19	179.47	96.47	124.34
150	6445.61	6632.65	13539.18	17527.54	179.58	150	166.74	105.12	247.55	166.04	179.58
200	7350.44	7601.17	15641.54	21028.11	233.99	200	232.49	130.31	317.47	232.01	233.99

The table reports the Ljung-Box statistic on trade durations (left side) and standardized residuals of trade duration (right side) for different stocks at different lags. The column of Lags is the number of lags we use to compute the statistic and the C\_Value is the critical value for the corresponding lags.

Table A.3: Estimated Results For Quote Durations

		MEO			MRK			RWE			TKA		
		Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat
Panel A	ln(dur(lag1))	0.309	0.004	<b>71.593</b>	0.301	0.006	<b>52.698</b>	0.301	0.004	<b>75.432</b>	0.208	0.005	<b>42.521</b>
	ln(dur(lag2))	0.041	0.004	<b>9.273</b>	0.128	0.004	<b>29.884</b>	0.129	0.004	<b>36.112</b>	0.019	0.004	<b>4.836</b>
	ln(dur(lag3))	0.166	0.003	<b>49.881</b>							0.026	0.004	<b>6.562</b>
	ln(dur(lag4))										-0.007	0.003	<b>-2.074</b>
	x/phi (lag1)	0.071	0.001	<b>57.240</b>	0.042	0.001	<b>30.717</b>	0.063	0.001	<b>51.971</b>	0.047	0.001	<b>42.333</b>
	x/phi (lag2)										0.020	0.001	<b>21.157</b>
	x/phi (lag3)										0.017	0.001	<b>16.951</b>
x/phi (lag4)										0.016	0.001	<b>23.269</b>	
cons	-0.551	0.016	<b>-35.430</b>	-0.402	0.012	<b>-32.464</b>	-1.104	0.013	<b>-83.048</b>	-1.804	0.018	<b>-102.047</b>	
Panel B	x/phi (lag1)*dummy	0.079	0.004	<b>21.456</b>	0.066	0.005	<b>13.058</b>	0.076	0.004	<b>19.017</b>	0.070	0.003	<b>24.953</b>
	trade_dur/expected_trade_dur	0.151	0.002	<b>69.252</b>	0.269	0.003	<b>107.500</b>	0.266	0.002	<b>108.070</b>	0.252	0.002	<b>104.887</b>
	lag_trade_dur/expected_trade_dur	0.047	0.003	<b>15.898</b>	0.007	0.002	<b>2.689</b>	0.025	0.003	<b>8.069</b>	0.089	0.003	<b>34.113</b>
	lag_expected_trade_dur	0.098	0.005	<b>20.691</b>	-0.067	0.004	<b>-18.465</b>	0.201	0.004	<b>46.054</b>	0.228	0.004	<b>51.473</b>
	delta spread	-11.509	0.179	<b>-64.331</b>	-12.433	0.182	<b>-68.396</b>	-13.404	0.115	<b>-116.456</b>	-8.812	0.089	<b>-98.486</b>
	lev.spread	-4.277	0.099	<b>-43.253</b>	-1.560	0.112	<b>-13.878</b>	-9.776	0.117	<b>-83.286</b>	-5.077	0.106	<b>-47.759</b>
	square root(vol)	-0.071	3.74E-04	<b>-190.596</b>	-0.066	0.001	<b>-126.347</b>	-0.040	1.79E-04	<b>-222.665</b>	-0.029	1.69E-04	<b>-170.251</b>
	abs(sign.vol)	9.71E-05	1.69E-06	<b>57.393</b>	-3.20E-05	2.75E-06	<b>-11.627</b>	2.78E-05	1.09E-06	<b>25.421</b>	2.98E-05	7.05E-07	<b>42.287</b>
	Back_quote_dur	4.62E-04	1.85E-04	<b>2.490</b>	0.004	1.15E-04	<b>33.731</b>	0.010	2.84E-04	<b>35.358</b>	0.001	3.05E-04	<b>2.122</b>
	$\Delta$ XLM	-0.001	0.000	<b>-1.777</b>	-0.006	0.000	<b>-14.094</b>	-0.005	0.001	<b>-7.949</b>	0.020	0.001	<b>19.092</b>
Panel C	Dummy1	0.180	0.011	<b>16.712</b>	-0.408	0.012	<b>-34.861</b>	0.529	0.009	<b>55.791</b>	0.179	0.010	<b>17.631</b>
	Dummy2	0.127	0.007	<b>18.559</b>	-0.591	0.009	<b>-66.455</b>	-0.025	0.006	<b>-4.223</b>	-0.253	0.007	<b>-35.328</b>
	Dummy3	-0.071	0.006	<b>-12.364</b>	-0.110	0.008	<b>-14.433</b>	-0.057	0.006	<b>-9.237</b>	-0.221	0.007	<b>-30.566</b>
	Dummy4	-0.114	0.006	<b>-17.959</b>	-0.137	0.009	<b>-15.168</b>	-0.006	0.007	<b>-0.966</b>	-0.265	0.010	<b>-27.096</b>
	Dummy5	-0.085	0.006	<b>-13.872</b>	-0.119	0.009	<b>-13.211</b>	0.020	0.007	<b>2.695</b>	-0.022	0.009	<b>-2.306</b>
	Dummy6	-0.290	0.007	<b>-41.148</b>	-0.338	0.008	<b>-45.045</b>	-0.048	0.007	<b>-6.443</b>	-0.202	0.011	<b>-18.196</b>
	Dummy7	0.045	0.005	<b>9.017</b>	0.111	0.008	<b>13.382</b>	-0.089	0.006	<b>-14.067</b>	0.080	0.008	<b>9.441</b>
	Dummy8	-0.212	0.005	<b>-43.645</b>	0.013	0.008	1.599	-0.248	0.005	<b>-45.427</b>	-0.266	0.007	<b>-35.762</b>

The table reports the estimated results for quote durations. Panel A presents the results for Log-ACD part, Panel B includes the results for the exogenous variables and Panel C consists of the results for dummy variables. Bold entries indicate the coefficients are significant at a 5% level.

Table A.4: Ljung-Box Statistics For Quote Durations and Deduced Standardized Residuals

Ljung-Box Statistics on Trade Durations						Ljung-Box Statistics on Corresponding Standardized Residuals					
Lags	MEO	MRK	RWE	TKA	C_Value	Lags	MEO	MRK	RWE	TKA	C_Value
5	120.094	34.113	297.782	492.920	11.070	5	9.61	6.28	8.21	7.93	11.07
10	134.796	42.798	369.755	611.703	18.307	10	14.16	16.12	9.96	50.26	18.31
15	164.762	52.638	392.250	687.889	24.996	15	24.68	23.33	11.12	73.81	25.00
20	181.157	63.902	424.918	727.079	31.410	20	46.48	33.47	12.14	117.30	31.41
50	278.866	117.195	650.198	952.415	67.505	50	76.23	48.55	52.44	256.63	67.50
100	408.867	169.944	1019.386	1373.740	124.342	100	152.76	92.94	128.93	350.57	124.34
150	559.907	212.972	1200.208	1623.935	179.581	150	216.94	129.18	176.10	446.76	179.58
200	664.119	257.827	1373.355	2121.540	233.994	200	276.92	177.21	214.65	522.38	233.99

The table reports the Ljung-Box statistic on quote durations (left side) and standardized residuals of quote duration (right side) for different stocks at different lags. The column of Lags is the number of lags we use to compute the statistic and the C\_Value is the critical value for the corresponding lags.



Table A.5: Estimated Results For Liquidity Activity Factor of  $XLM^{5000}$

		MEO			MRK			RWE			TKA		
		Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat
Panel A	$g(t-1)$	0.912	0.017	<b>54.593</b>	0.902	0.024	<b>38.165</b>	0.790	0.052	<b>15.224</b>	0.854	0.026	<b>32.492</b>
	$activity(t-1)$	1.071	0.044	<b>24.088</b>	0.991	0.051	<b>19.485</b>	1.080	0.034	<b>31.700</b>	0.865	0.031	<b>27.978</b>
	$activity(t-2)$	-0.862	0.057	<b>-15.133</b>	-0.779	0.068	<b>-11.417</b>	-0.740	0.082	<b>-9.041</b>	-0.619	0.049	<b>-12.703</b>
	cons	-1.256	0.139	<b>-9.051</b>	-1.239	0.148	<b>-8.356</b>	-1.919	0.096	<b>-20.042</b>	-1.541	0.077	<b>-19.912</b>
Panel B	Expect Trade Dur	0.118	0.040	<b>2.923</b>	0.162	0.037	<b>4.407</b>	0.174	0.030	<b>5.824</b>	0.174	0.024	<b>7.140</b>
	Expect Quote Dur	-0.065	0.024	<b>-2.711</b>	-0.024	0.021	-1.157	-0.120	0.022	<b>-5.488</b>	-0.033	0.014	<b>-2.408</b>
	delta spread	2.881	0.757	<b>3.806</b>	2.582	1.028	<b>2.511</b>	4.049	0.788	<b>5.137</b>	2.154	0.594	<b>3.628</b>
	lev.spread	3.612	1.128	<b>3.201</b>	2.899	1.576	1.840	9.978	1.326	<b>7.524</b>	8.295	0.898	<b>9.238</b>
	square root(vol)	-0.014	0.003	<b>-4.801</b>	-0.020	0.004	<b>-5.364</b>	-0.004	0.002	<b>-2.635</b>	-0.001	0.001	-1.313
	abs(sign.vol)	-6.18E-06	1.94E-05	-0.319	-5.62E-06	3.10E-05	-0.181	-1.76E-05	0.000	-1.808	-1.86E-05	4.52E-06	<b>-4.107</b>
	Back_quote_dur $\Delta XLM$	0.004 0.009	7.94E-04 0.003	<b>5.614</b> <b>2.587</b>	0.003 0.007	6.40E-04 0.003	<b>4.146</b> <b>2.250</b>	0.007 0.011	1.10E-03 0.003	<b>6.679</b> <b>3.339</b>	0.003 0.036	1.03E-03 0.006	<b>2.800</b> <b>6.027</b>
Panel C	Dummy1	0.006	0.143	0.042	0.370	0.142	<b>2.612</b>	-0.002	0.088	-0.028	0.140	0.072	1.949
	Dummy2	-0.226	0.092	<b>-2.463</b>	0.168	0.092	1.830	-0.023	0.065	-0.349	0.116	0.054	<b>2.139</b>
	Dummy3	-0.130	0.089	-1.459	0.147	0.095	1.548	-0.027	0.065	-0.420	0.143	0.057	<b>2.508</b>
	Dummy4	-0.176	0.103	-1.719	-0.045	0.098	-0.454	-0.065	0.067	-0.967	0.087	0.066	1.315
	Dummy5	-0.137	0.095	-1.437	0.202	0.115	1.754	-0.124	0.084	-1.466	-0.042	0.075	-0.558
	Dummy6	-0.127	0.100	-1.261	0.009	0.099	0.095	-0.097	0.078	-1.246	-0.019	0.075	-0.258
	Dummy7	-0.165	0.077	<b>-2.128</b>	-0.155	0.093	-1.665	-0.109	0.070	-1.551	-0.008	0.066	-0.119
	Dummy8	-0.066	0.067	-0.989	0.060	0.089	0.675	0.034	0.055	0.613	0.149	0.052	<b>2.854</b>

The table reports the estimated results for Activity factor. Panel A presents the results for GLARMA part, Panel B includes the results for the exogenous variables and Panel C consists of the results for dummy variables. Bold entries indicate the coefficients are significant at a 5% level.

Table A.6: Ljung-Box Statistics For Activity Factor of  $XLM^{5000}$  and Deduced Standardized Residuals

Ljung-Box Statistics on Trade Durations						Ljung-Box Statistics on Corresponding Standardized Residuals					
Lags	MEO	MRK	RWE	TKA	C_Value	Lags	MEO	MRK	RWE	TKA	C_Value
5	1403.77	963.02	2023.81	2004.82	11.07	5	10.62	9.70	4.86	5.52	11.07
10	1617.65	1144.00	2273.72	2420.69	18.31	10	13.94	15.39	6.36	13.95	18.31
15	1774.37	1293.30	2385.44	2641.78	25.00	15	17.93	21.91	11.19	21.31	25.00
20	1986.75	1391.13	2466.59	2782.23	31.41	20	23.94	25.64	13.45	28.13	31.41
50	2424.58	1694.06	2774.86	3205.48	67.50	50	71.28	49.87	46.35	64.74	67.50
100	2690.33	1890.80	3068.62	3545.26	124.34	100	127.19	86.74	96.45	123.66	124.34
150	2883.88	1962.55	3220.05	3769.42	179.58	150	186.08	131.70	155.26	176.11	179.58
200	2949.67	1998.66	3272.72	3943.11	233.99	200	246.23	172.25	201.76	229.07	233.99

The table reports the Ljung-Box statistic on activity factor (left side) and standardized residuals of activity factor (right side) for different stocks at different lags. The column of Lags is the number of lags we use to compute the statistic and the C\_Value is the critical value for the corresponding lags.

Table A.7: Estimated Results For Direction Factor of  $XLM^{5000}$

		MEO			MRK			RWE			TKA		
		Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat
Panel A	$g(t-1)$	0.64	0.05	<b>12.73</b>	0.54	0.05	<b>11.28</b>	0.44	0.03	<b>17.23</b>	0.48	0.03	<b>18.18</b>
	$g(t-2)$	-0.04	0.05	-0.91	0.03	0.04	0.74	0.22	0.02	<b>9.39</b>	0.09	0.02	<b>3.76</b>
	$g(t-3)$												
	$direction(t-1)$	-0.71	0.04	<b>-17.92</b>	-0.81	0.04	<b>-18.47</b>	-1.39	0.04	<b>-32.70</b>	-1.13	0.04	<b>-30.67</b>
	cons	-0.06	0.09	-0.68	-0.08	0.12	-0.68	0.20	0.10	1.92	0.23	0.09	<b>2.72</b>
Panel B	Expect Quote	0.00	0.01	-0.41	0.02	0.01	<b>2.04</b>	-0.01	0.01	-1.44	0.00	0.01	-0.14
	delta spread	3.67	0.84	<b>4.36</b>	4.47	1.01	<b>4.41</b>	8.29	1.08	<b>7.67</b>	5.86	0.74	<b>7.96</b>
	lev.spread	-1.75	1.20	-1.47	-1.13	1.73	-0.65	-6.34	1.85	<b>-3.43</b>	-5.87	1.06	<b>-5.51</b>
	square root(vol)	2.79E-03	3.65E-03	0.76	4.22E-03	0.00	0.85	2.25E-03	3.05E-03	0.74	2.68E-04	1.68E-03	0.16
	abs(sign.vol)	2.58E-05	2.33E-05	1.10	2.31E-05	4.13E-05	0.56	-1.01E-05	1.63E-05	-0.62	-8.46E-06	6.91E-06	-1.22
	Back_quote_dur	-4.54E-03	8.62E-04	<b>-5.27</b>	-2.50E-03	6.77E-04	<b>-3.69</b>	-0.01	1.54E-03	<b>-4.94</b>	-7.06E-03	1.40E-03	<b>-5.06</b>
	$\Delta XLM$	-0.05	0.01	<b>-9.00</b>	-0.02	0.00	<b>-5.39</b>	0.00	0.00	-1.03	-0.06	0.01	<b>-6.36</b>
Panel C	Dummy1	0.27	0.16	1.69	-0.02	0.16	-0.12	0.22	0.14	1.60	0.16	0.10	1.49
	Dummy2	0.11	0.11	0.97	0.06	0.11	0.54	0.10	0.11	0.92	0.01	0.09	0.10
	Dummy3	0.04	0.11	0.38	0.08	0.12	0.68	0.07	0.11	0.68	-0.03	0.09	-0.38
	Dummy4	-0.03	0.12	-0.25	0.03	0.13	0.23	0.01	0.12	0.12	-0.22	0.10	<b>-2.28</b>
	Dummy5	0.12	0.11	1.09	-0.01	0.13	-0.06	0.21	0.13	1.62	0.03	0.11	0.23
	Dummy6	0.10	0.11	0.93	0.20	0.13	1.57	0.12	0.13	0.94	-0.13	0.11	-1.12
	Dummy7	0.11	0.10	1.15	0.06	0.13	0.48	0.12	0.12	1.04	0.01	0.10	0.14
	Dummy8	0.11	0.08	1.29	-0.06	0.11	-0.52	-0.07	0.09	-0.72	0.01	0.08	0.11

The table reports the estimated results for Direction factor. Panel A presents the results for GLARMA part, Panel B includes the results for the exogenous variables and Panel C consists of the results for dummy variables. Bold entries indicate the coefficients are significant at a 5% level.

Table A.8: Ljung-Box Statistics For Direction Factor of  $XLM^{5000}$  and Deduced Standardized Residuals

Ljung-Box Statistics on Trade Durations						Ljung-Box Statistics on Corresponding Standardized Residuals					
Lags	MEO	MRK	RWE	TKA	C_Value	Lags	MEO	MRK	RWE	TKA	C_Value
5	936.12	753.51	2248.85	2436.79	11.07	5	2.97	6.02	7.18	11.98	11.07
10	938.76	761.46	2416.30	2474.64	18.31	10	5.85	12.58	23.13	27.01	18.31
15	942.35	762.92	2429.77	2486.77	25.00	15	9.68	14.56	24.41	33.03	25.00
20	945.28	767.59	2435.74	2505.57	31.41	20	10.98	17.77	28.31	49.31	31.41
50	983.90	798.82	2489.91	2528.93	67.50	50	45.43	46.71	52.98	75.56	67.50
100	1050.19	880.95	2595.57	2619.30	124.34	100	88.78	105.35	116.60	121.02	124.34
150	1124.15	949.12	2648.78	2703.59	179.58	150	154.47	150.86	151.78	168.37	179.58
200	1181.43	1004.79	2775.72	2759.53	233.99	200	193.91	188.65	202.96	212.86	233.99

The table reports the Ljung-Box statistic on direction factor (left side) and standardized residuals of direction factor (right side) for different stocks at different lags. The column of Lags is the number of lags we use to compute the statistic and the C\_Value is the critical value for the corresponding lags.

Table A.9: Estimated Results For Liquidity Size Factor of  $XLM^{5000}$

		MEO			MRK			RWE			TKA		
		Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat	Est	STD	T-Stat
Panel A	$g(t-1)$	0.94	0.01	<b>146.73</b>	0.92	0.01	<b>135.73</b>	0.75	0.02	<b>34.85</b>	-0.22	0.07	<b>-3.11</b>
	$g(t-2)$	-0.06	0.00	<b>-21.67</b>	-0.04	0.00	<b>-21.67</b>	0.45	0.03	<b>15.93</b>	0.79	0.04	<b>21.76</b>
	$g(t-3)$							-0.23	0.02	<b>-10.69</b>	0.10	0.04	<b>2.75</b>
	$Size(t-1)$	0.05	0.00	<b>16.63</b>	0.03	0.00	<b>15.64</b>	-0.09	0.00	<b>-37.45</b>	-0.17	0.01	<b>-31.05</b>
	$Size(t-2)$							0.08	0.00	<b>34.27</b>	-0.04	0.01	<b>-3.73</b>
	$Size(t-3)$										0.11	0.01	<b>11.27</b>
	cons	-0.74	0.06	<b>-12.31</b>	-1.15	0.06	<b>-18.89</b>	-0.44	0.05	<b>-9.60</b>	0.05	0.05	1.15
Panel B	Expect QuoteQuote Dur	0.00	0.00	-0.92	0.00	0.00	-1.62	0.00	0.00	-0.69	0.00	0.00	-1.59
	Direction(t)	-0.05	0.01	<b>-3.26</b>	-0.03	0.01	<b>-2.08</b>	-0.02	0.01	-1.58	-0.11	0.01	<b>-8.87</b>
	Direction(t-1)	0.02	0.02	1.09	0.03	0.02	1.72	0.16	0.01	<b>12.72</b>	0.10	0.02	<b>5.53</b>
	delta spread	1.44	0.43	<b>3.33</b>	1.70	0.48	<b>3.53</b>	3.64	0.40	<b>9.18</b>	1.78	0.35	<b>5.04</b>
	lev.spread	0.90	0.59	1.52	0.97	0.70	1.38	1.91	0.70	<b>2.72</b>	1.83	0.38	<b>4.79</b>
	square root(vol)	0.01	0.00	<b>4.53</b>	0.01	0.00	<b>7.08</b>	0.01	0.00	<b>13.89</b>	0.01	0.00	<b>7.42</b>
	abs(sign.vol)	0.00	0.00	0.20	0.00	0.00	0.30	0.00	0.00	-1.82	0.00	0.00	-1.35
	Back_quote_dur	0.00	0.00	<b>7.80</b>	0.00	0.00	<b>6.37</b>	0.00	0.00	<b>6.25</b>	0.00	0.00	<b>4.22</b>
$\Delta XLM$	-0.01	0.00	-1.93	0.00	0.00	-1.13	-0.01	0.00	<b>-5.25</b>	-0.03	0.01	<b>-5.87</b>	
Panel C	Dummy1	-0.18	0.08	<b>-2.42</b>	-0.07	0.06	-1.16	-0.32	0.05	<b>-6.29</b>	-0.29	0.05	<b>-5.79</b>
	Dummy2	-0.06	0.06	-1.14	-0.07	0.05	-1.38	0.05	0.04	1.21	-0.13	0.04	<b>-3.05</b>
	Dummy3	0.06	0.05	1.09	-0.05	0.05	-0.85	-0.12	0.04	<b>-3.03</b>	-0.11	0.04	<b>-2.46</b>
	Dummy4	-0.15	0.06	<b>-2.59</b>	0.18	0.06	<b>3.22</b>	-0.09	0.04	<b>-2.17</b>	-0.21	0.05	<b>-4.31</b>
	Dummy5	-0.16	0.06	<b>-2.76</b>	-0.06	0.06	-0.96	-0.23	0.05	<b>-4.74</b>	-0.06	0.06	-1.10
	Dummy6	-0.13	0.06	<b>-2.15</b>	0.19	0.05	<b>3.63</b>	-0.15	0.04	<b>-3.28</b>	-0.09	0.06	-1.56
	Dummy7	-0.04	0.05	-0.88	0.21	0.05	<b>3.91</b>	0.03	0.04	0.84	-0.13	0.05	<b>-2.72</b>
	Dummy8	-0.06	0.04	-1.52	0.04	0.05	0.89	0.00	0.04	-0.12	-0.07	0.04	-1.72

The table reports the estimated results for Size factor. Panel A presents the results for GLARMA part, Panel B includes the results for the exogenous variables and Panel C consists of the results for dummy variables. Bold entries indicate the coefficients are significant at a 5% level.

Table A.10: Ljung-Box Statistics For Size Factor of  $XLM^{5000}$  and Deduced Standardized Residuals

Ljung-Box Statistics on Trade Durations						Ljung-Box Statistics on Corresponding Standardized Residuals					
Lags	MEO	MRK	RWE	TKA	C_Value	Lags	MEO	MRK	RWE	TKA	C_Value
5	2209.94	1563.99	2558.04	2290.37	11.07	5	99.40	83.52	76.66	55.92	11.07
10	2843.94	2192.57	2981.30	2644.72	18.31	10	109.68	112.84	90.15	60.68	18.31
15	3344.19	2611.31	3409.76	2747.69	25.00	15	112.38	127.07	99.77	61.36	25.00
20	3810.83	2947.08	3791.73	2800.32	31.41	20	113.32	136.01	107.60	66.46	31.41
50	5921.45	4073.51	5031.22	2898.91	67.50	50	146.46	165.35	142.29	91.11	67.50
100	7594.86	4862.05	6008.86	2981.64	124.34	100	197.35	237.50	184.51	149.92	124.34
150	9426.42	5011.23	6671.05	3053.38	179.58	150	294.90	292.40	268.07	215.81	179.58
200	10904.26	5257.11	6960.31	3151.29	233.99	200	333.15	339.92	317.39	272.71	233.99

The table reports the Ljung-Box statistic on size factor (left side) and standardized residuals of size factor (right side) for different stocks at different lags. The column of Lags is the number of lags we use to compute the statistic and the C\_Value is the critical value for the corresponding lags.