Investigating the Relationship between Trade Duration and Liquidity: Evidence from China

Mingyue Qi
Institut Mines-Telecom/Telecom Ecole de Management, France
Email: mingyue.qi@telecom-em.eu

Bin Li
University of Chinese Academy of Sciences/School of Management, Beijing, China
Email: libincas@126.com

Abstract—This paper investigates the relationship between trade duration and liquidity of Chinese stock market. By using data of ten stocks in Chinese stock market, we employ a Weibull ACD model to decompose trade duration into two components: the expected and the unexpected duration. Then we analyze whether trade duration affects liquidity with regressions. Finally, we find that there exists a strong dependence between consecutive durations especially for liquid stocks. Both the expected and unexpected duration could explain the variation of bid-ask spread but the evidence is mixed in the depth equation. The unexpected duration contributes more to the change in liquidity than the expected duration.

Index Terms—trade duration, liquidity, spread, depth, WACD model

I. INTRODUCTION

Trade duration, which refers to the waiting time between two consecutive transactions of equity, is the main concept of the autoregressive conditional duration (ACD) model [1]. This model is aim to investigate the durations between events such as: trades, quotes, price changes with consideration of the data which are spaced irregularly in time [2]. We use the distributions such as the Exponential, Gamma and Weibull to model ACD structures [3]. The Weibull distribution is more flexible and therefore plays an important role in ACD modelling [4].

Stock liquidity has been an important issue in financial market. It is influenced by trading activity from several perspectives. The relationship between liquidity and trading activity has drawn lots of attentions. Easley and O’Hara [5] prove that trades convey information which causes price changes and liquidity variations. Biais et al. [6] study the interaction between the order book and order flow. Using data from a pure limit order market, they analyze the supply and demand of liquidity, and they find that after large sales (purchases), there is often a new sell (buy) order placed within the quotes to provide liquidity. Hendershott et al. [7] provide the first analysis on whether algorithmic (AT) trading increase market liquidity. They find AT narrows spreads, reduces adverse selection, and reduces trade-related price discovery for NYSE stocks, particular the larger stocks, indicating that AT improves liquidity. Chordia et al. [8] document positive feedback exists between liquidity and trading activity. Daniel, Robert and Philip [9] analyzes the relationship between liquidity and trading characters by adopting the six liquidity proxies [10], Pastor and Stambaugh [11], Liu [12]: bid–ask spread, stock turnover, the illiquidity ratio, the return reversal measure, the zero return measure (proportion of zero daily returns), and turnover adjusted number of zero daily volumes. Stoll [3] and Chordia et al. [13], [14] analyzes how the stock trading characteristics influence liquidity from the following perspectives: stock price, trading volume and volatility without the trading duration. They find a secular downtrend spread and an upward trend depth and trading volume. They also note strong day-of-the-week effects that Friday accompanies a significant decrease in trading activity and liquidity. It indicates the relevance between duration and liquidity.

However, the relationship between trade duration and liquidity has not been explored by previous studies. To investigate the relationship between trade duration and liquidity could broaden the measurement of liquidity, and also help investors to make the right decision according to different conditions.

In this paper, we investigate the relationship with high frequency data from Chinese stock market. We decompose the trade duration into two components with a Weibull autoregressive conditional duration (WACD) model. Using quoted bid-ask spread and depth as liquidity proxies, we propose two regressions to analyze how duration changes spread and depth. Unlike previous literatures, the decomposition of trade duration in this paper allows us to study the information contained in expected and unexpected duration and their role in affecting liquidity.

The reminder of the paper is structured as follows: Section II introduces the background of Chinese stock market. Section III describes our liquidity measures and data preparation. Section IV introduces the autoregressive conditional duration model and its extension model. Section V presents the estimation results of trade durations; we also analyze their effects on stock liquidity in this section. Section 6 concludes our work.

©2013 Engineering and Technology Publishing
doi: 10.12720/joams.1.3.332-337
II. THE BACKGROUND OF CHINESE STOCK MARKET

Chinese stock market is organized as a fully automated, pure order-driven market. It consists of two stock exchanges: the Shanghai exchange and the Shenzhen exchange. Both stock exchanges are open five days a week from 9:30 a.m. to 3:00 p.m. with a trading break from 11:30 a.m. to 1:00 p.m. daily. For each trading day, there is a pre-trading call auction period from 9:15 to 9:25 before continuous auction trading. During the continuous trading session, orders are submitted, modified or cancelled. A trade takes place whenever a counterpart order hits the quotes. The exchange automatically matches all the orders according to strict price and time priority. However, an order may lose priority if modified. Stocks are quoted in RMBs. The minimum price variation (tick) equals 0.01. The minimum trade size is 100 shares. The daily price limit is 10 percent of the previous day’s closing price. Chan et al. [15] documents that such price limit mechanisms are employed in many stock exchanges around the world, including Austria, Belgium, France, Italy, Japan and Korea.

There are two main share types in China, the A Shares and the B Shares. An issuing firm can have one, or both share types. The investors of the A Shares and the B Shares are quite different. Until early 2001, the B Shares were restricted to foreign investors, typically foreign institutions, while the A Shares were restricted to domestic investors, and became available to select foreign investors in 2005.

The A shares are considered to represent the Chinese stock market because A Shares clearly dominate the market in every aspects, such as the number of shares issued, market capitalization, and trading activity. In this paper, we only focus on studying the relationship of duration and liquidity of the A shares stocks.

III. LIQUIDITY MEASURES AND DATA PREPARATION

A. Liquidity Measures

It is widely recognized that the concept of market liquidity cannot be captured by a single measure. Kyle [16] identifies three main dimensions of liquidity: tightness, depth and resiliency. First of all, tightness refers to the narrowness of the bid-ask spread which is often measured by the quote-based bid-ask spread. It symbols the transaction costs for traders. Second, depth is the market’s ability to absorb and execute large orders with minimal price impact, which is often measured by the quoted depth. At last, resiliency is the speed with which prices recover from a random, uninformative shock.

Amihud and Mendelson [17] prove that there is evidence that bid-ask spread has some relevance with asset pricing. Engle and Lange [18] realize the importance of depth, and they propose a new intraday measure of market depth. Easley et al. [19] highlight the importance of both bid-ask spread and depth. The resiliency dimension of liquidity is difficult to measure accurately. Large [20] measures the resiliency by viewing orders and cancellations as a mutually-exciting ten-variate Hawkes point processes. There are alternative ways to derive liquidity measures from daily, monthly or even annual return and volume data, Goyenko et al. [21] provide a detailed comparison of commonly used low frequency liquidity proxies, they find that most monthly and annual low-frequency measures capture high-frequency measures of transaction costs.

Since our aim is to investigate the high frequency dynamics between trade duration and liquidity, we use two high frequency liquidity measures, the quoted bid-ask spread and the quoted depth, to proceed with our research. We decide to eliminate the dimension of resiliency because it lacks of a widely accepted accurate measure.

B. Data Preparation

The sample of this study is 10 randomly selected A shares stocks of Chinese stock market. To be included in our sample, a stock has to meet the following criteria: First, during the sample period, the stock should not be under “special treatment (ST)”. If a stock is being special treated by its stock exchange, the daily price limit will be the half of the regularly traded stock.

Shenoy and Zhang [22] find that the difference of price limits may affect the process of regular trading which may introduce some bias. Second, the stock has to be listed before December 2010. This criterion ensures that the selected stock has no less than half a year trading history, which avoids the bias caused by IPOs.

According to the criteria set above, our sample stocks and their codes are: LHGF (000532), ZXGA (000839), KHSW (002022), LDGD (002189), SZTY (300002), JRXF (300116), FXXC (300139), RZG (600017), ZFSY (600058) and BHGF (601678).

The data are extracted from CSMAR database from September 2011 to June 2012. The CSMAR database is a collection of all trades in both Shanghai and Shenzhen stock exchange. Each trade is recorded with following fields: the trading time, volume, trade direction (the trade is buyer-or seller-initiated), five best bid and ask quote prices and their corresponding volume. To avoid wrong records, we filter the data based on the following principals: firstly, since we only investigate the relationship between trading duration and liquidity of the continuous trading period, we delete all the trade records
before 9:30 a.m. and after 15:00 p.m. secondly, the records with quote price, transaction price or bid-ask spread of zero are deleted. Thirdly, the trading days when the stock price reached its price limit are deleted because the price discovery process of a stock is limited by its daily price limits and the unchanged price will cause errors when estimating our model. Table I reports summary statistics of our data. ZXGA is the most frequently traded stock with an average of about 1600 transactions a day while JRXF has the least transactions on average, and so on.

IV. ECONOMIC METHODOLOGY

Traditional econometric models which are based on fixed interval analysis fail to work since transaction data arrives in irregular time intervals, while traditional econometric models are based on fixed interval analysis. Engle and Russell [1] firstly model the financial duration using a point process and their model is called the ACD (Autoregressive Conditional Duration) model.

The ACD model assumes a stochastic process which is

\[ x(t) = \epsilon_t \]

\[ \epsilon_t \sim iid \]

Assume it is the conditional probability \( x(t) \) given the duration between \( t_{i-1} \) and \( t_i \). Then the expected duration is the conditional intensity of the process is,

\[ \lambda(t | x_{t_{i-1}}, \ldots, x_t) = \left( \Gamma(1 + \gamma) \psi_N \right)^{1-1} \gamma (t - t_{t_{i-1}})^{ \gamma - 1} \] (5)

where \( \Gamma() \) is the gamma function, \( \gamma \) is the Weibull parameter. The conditional intensity is now a two parameter family which can exhibit either increasing or decreasing hazard functions. This predicts especially long durations more or less likely than for the exponential depending on whether \( \gamma \) is less or greater than 1 respectively. The log likelihood for the Weibull ACD (WACD) is,

\[ \log L = \sum_{j=1}^{N(T)} \ln \left( \frac{1}{\lambda} \right) + \gamma \ln \Gamma(1 + \gamma) x_j^{\gamma - 1} - \Gamma(1 + \gamma) x_j^{\gamma - 1} \] (6)

For some parameterization \( \psi \). When \( \gamma = 1 \), the log likelihood reduces to the log likelihood for the exponential ACD. The parameters can be easily obtained by maximum-likelihood estimation.

High frequency intra-day data is commonly found to be affected by the time of the trading day, such as return and volume. Duration is no exception, trading tends to be more intensive at opening and closing of the market than during the course of the day. We firstly adjust the original duration, \( x_t \), for intraday effects to obtain the diurnally adjusted duration, \( \tilde{x}_t \), which is given by \( \tilde{x}_t = x_t / \phi(x_t, \theta) \). Then the expected duration is and \( \theta \) can be \( \hat{E}(x_t | \phi(x_t, \theta)) \) jointly estimated using maximum likelihood method. As Engle and Russell [1] point out, joint estimation causes a heavy computational burden, they turn to a two-stage estimation procedure. They first remove the intraday effects of duration with a cubic spline function, and then they model the adjusted duration with ACD. The results of two-stage estimation are proved to be quite similar to the joint estimation results, yet it reduces the computational burden greatly. In our paper, we adopt this method and develop it with the expected and unexpected duration.

V. EMPIRICAL RESULTS

A. The WACD Result

We fit the diurnally adjusted duration in a WACD (p, q) model using the two-stage maximum likelihood method. The Akaike Information Criterion (AIC) is employed to determine the optimal lag length. In this paper, all the sample stocks have the same optimal WACD lag length of one which indicates the WACD (1, 1) fits the data best. The estimation results are presented in Table II.

For all the stocks, \( \alpha \) and \( \beta \) are statistically significant, indicating a close interdependence in the trade duration series. The positive coefficients \( \alpha \) and \( \beta \) suggests a high degree of dependence of consecutive durations. They provide a strong evidence that the arrival rates of past
trades can determine the time when the next trade occurs, implying that a short duration is likely to be followed by a short duration while a long duration followed by a long one. The current trade duration is useful in determining the future duration which, to some extent, could explain the clustering of trades at certain time of the day.

The Weibull parameter \( \gamma \) of each stock is significantly greater than 1, indicating an upward sloping hazard and the positive dependence in the duration. Engle and Russell [1] analyze the trade data of U.S. stock market; they find the opposite results that all \( \gamma \) are less than 1. We attribute the difference to the inactivity of Chinese stock market. Comparing with the developed financial markets, Chinese stocks have longer trade durations.

The most frequently traded stocks such as ZXGA, LHGF and KHSW, their sum of \( \alpha \) and \( \beta \) is higher than other stocks, implying the existence of the most persistent effect of their past trading frequency on current trade’s arrival time. In contrast, JRXF, the least traded stock has the lowest sum of \( \alpha \), \( \beta \), which indicates weakest dependence between consecutive durations. Comparison across stocks show evidence that more frequently traded stocks appears to have cluster in trades. One possible explanation is that the institutional investors favor high liquid stocks. They may possess an information advantage and are unwilling to cause a large price impact. In order to profit with the advantage, they have to split a large order into a few smaller ones and place them one by one in the limit order book to avoid the unfavorable price impact.

<table>
<thead>
<tr>
<th>Stock</th>
<th>( \omega )</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \alpha + \beta )</th>
<th>( \gamma )</th>
<th>Likelihood Function Value</th>
<th>AIC</th>
<th>Ljung-Box Q(5)</th>
<th>Ljung-Box Q(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LHGF</td>
<td>0.038*** (0.000)</td>
<td>0.081*** (0.001)</td>
<td>0.883*** (0.002)</td>
<td>0.964</td>
<td>1.393*** (0.002)</td>
<td>-256242.96</td>
<td>5.37</td>
<td>20.740 (0.654)</td>
<td>15.170 (0.508)</td>
</tr>
<tr>
<td>ZXGA</td>
<td>0.055*** (0.001)</td>
<td>0.100*** (0.001)</td>
<td>0.847*** (0.002)</td>
<td>0.947</td>
<td>1.400*** (0.002)</td>
<td>-263212.12</td>
<td>5.36</td>
<td>18.660 (0.770)</td>
<td>33.580 (0.584)</td>
</tr>
<tr>
<td>KHSW</td>
<td>0.038*** (0.001)</td>
<td>0.071*** (0.001)</td>
<td>0.893*** (0.002)</td>
<td>0.964</td>
<td>1.370*** (0.002)</td>
<td>-265763.92</td>
<td>5.83</td>
<td>28.950 (0.222)</td>
<td>36.770 (0.433)</td>
</tr>
<tr>
<td>LDGD</td>
<td>0.086*** (0.004)</td>
<td>0.103*** (0.002)</td>
<td>0.814*** (0.005)</td>
<td>0.914</td>
<td>1.111*** (0.003)</td>
<td>-98388.52</td>
<td>5.81</td>
<td>30.410 (0.171)</td>
<td>44.190 (0.164)</td>
</tr>
<tr>
<td>SZTY</td>
<td>0.046*** (0.001)</td>
<td>0.084*** (0.001)</td>
<td>0.873*** (0.002)</td>
<td>0.957</td>
<td>1.334*** (0.002)</td>
<td>-233157.33</td>
<td>5.44</td>
<td>29.270 (0.210)</td>
<td>39.320 (0.323)</td>
</tr>
<tr>
<td>JRXF</td>
<td>0.196*** (0.009)</td>
<td>0.128*** (0.004)</td>
<td>0.681*** (0.012)</td>
<td>0.809</td>
<td>1.026*** (0.003)</td>
<td>-64128.09</td>
<td>5.98</td>
<td>19.130 (0.745)</td>
<td>31.480 (0.493)</td>
</tr>
<tr>
<td>FXXC</td>
<td>0.112*** (0.004)</td>
<td>0.131*** (0.003)</td>
<td>0.758*** (0.006)</td>
<td>0.891</td>
<td>1.070*** (0.002)</td>
<td>-102039.43</td>
<td>5.98</td>
<td>12.350 (0.976)</td>
<td>23.280 (0.950)</td>
</tr>
<tr>
<td>RZG</td>
<td>0.082*** (0.003)</td>
<td>0.070*** (0.002)</td>
<td>0.850*** (0.004)</td>
<td>0.920</td>
<td>1.315*** (0.003)</td>
<td>-131766.27</td>
<td>5.42</td>
<td>9.840 (0.995)</td>
<td>27.120 (0.857)</td>
</tr>
<tr>
<td>ZFSY</td>
<td>0.066*** (0.002)</td>
<td>0.091*** (0.001)</td>
<td>0.846*** (0.003)</td>
<td>0.937</td>
<td>1.522*** (0.002)</td>
<td>-206518.98</td>
<td>5.16</td>
<td>27.310 (0.073)</td>
<td>40.160 (0.021)</td>
</tr>
<tr>
<td>LHGF</td>
<td>0.051*** (0.002)</td>
<td>0.077*** (0.001)</td>
<td>0.874*** (0.002)</td>
<td>0.951</td>
<td>1.508*** (0.002)</td>
<td>-193060.64</td>
<td>5.18</td>
<td>29.590 (0.042)</td>
<td>43.960 (0.011)</td>
</tr>
</tbody>
</table>

Note: The p-values are in the parentheses, Symbols ***, **, * means the coefficient is significant at 1%, 5%, 10% level respectively.

We use the the Ljung–Box Q-statistics to test autocorrelation in residuals. For almost all the stocks, at a lag of 5 and 10, we fail to reject the null hypothesis that the residual series are auto correlated, which suggests our WACD (1, 1) model is specified correctly.

B. The Regression Result

The analysis above implies that some traders favour more liquid stocks, thus resulting higher persistence of trade duration. As stated in section III, modelling of the duration in an ACD framework allows a decomposition of the conditional duration into two components, the expected (conditional) and the unexpected (standard) duration. In this subsection, we proceed to investigate which component of the trade duration, the expected duration or the unexpected duration, has more influence on liquidity.

Quoted bid-ask spread and depth as liquidity proxies are modelled as follows,

\[
\text{Spread}_t = \mu + \lambda_s \cdot \text{Spread}_{t-1} + \beta_s \cdot \text{ExpectS}_t + \beta_d \cdot \text{UnexpectS}_t + \epsilon_t, \quad (7)
\]

\[
\text{Depth}_t = \mu + \lambda_d \cdot \text{Depth}_{t-1} + \beta_s \cdot \text{ExpectD}_t + \beta_d \cdot \text{UnexpectD}_t + \xi_t \quad (8)
\]

In equation (7) and (8), Spread\(_t\) and Depth\(_t\) are the quoted bid-ask spread and depth at time \( t \) respectively, we add the lagged spread and depth into the corresponding equation to eliminate serial autocorrelation. ExpectS\(_t\) and ExpectD\(_t\) are the expected durations at time \( t \) which is calculated using the WACD model in the above subsection. Similarly, UnexpectS\(_t\) and UnexpectD\(_t\) are the unexpected durations at time \( t \). \( \epsilon_t \) and \( \xi_t \) denote the i.i.d. errors. Regression results are presented in Table III. Since the lagged spread and depth are used to control serial autocorrelation, their coefficients are not shown.
between consecutive durations is stronger for higher liquid following conclusions: first of all, the dependence affects the liquidity. Using two regressions, we reach the investigating how information contained in trade duration and liquidity of Chinese stock market. Unlike stocks have better tightness. This is consistent with Biais [6] which find that, in a pure limit order book market, traders consume liquidity when it’s abundant and replenish the spread when it’s too large. The traders’ behavior makes the spread revert to its mean. Compared to the expected duration, the unexpected duration’s coefficient makes the spread wider.

All the coefficients $\beta_1$ and $\beta_2$, are significantly negative in Table III, which suggests long trade duration always lead to narrow spread and infrequently traded stocks have better tightness. This is consistent with Biais et al. [6] which find that, in a pure limit order book market, traders consume liquidity when it’s abundant and replenish the spread when it’s too large. The traders' behavior makes the spread revert to its mean. Compared to the expected duration, the unexpected duration’s coefficient has bigger t value which implies that the unexpected duration plays a more important role in determining the liquidity. This is consistent with Engle and Lange [19] who document that the change in intra-day liquidity is attributed more to the unexpected duration than the expected duration. The ACD model assumes the information from past durations are captured by the expected duration while additional information beyond that can be inferred from the past trading frequency is included in the unexpected duration. The evidence in the depth equation is mixed but most coefficients of the depth equation are significantly positive. Long trade duration gives the traders enough time to

<table>
<thead>
<tr>
<th>Equation (7)</th>
<th>$\beta_{1,S}$</th>
<th>$\beta_{2,S}$</th>
<th>$R^2$</th>
<th>Durbin-Watson test</th>
<th>Equation (8)</th>
<th>$\beta_{1,D}$</th>
<th>$\beta_{2,D}$</th>
<th>$R^2$</th>
<th>Durbin-Watson test</th>
<th>Number of Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>LHGF</td>
<td>-0.039***(-7.071)</td>
<td>-0.040***(-2.286)</td>
<td>0.17</td>
<td>2.14</td>
<td>LHGF</td>
<td>0.003***(+1.217)</td>
<td>0.005***(+8.334)</td>
<td>0.92</td>
<td>2.29</td>
<td>279449</td>
</tr>
<tr>
<td>ZGGA</td>
<td>-0.034***(-11.168)</td>
<td>-0.030***(-3.141)</td>
<td>0.35</td>
<td>2.19</td>
<td>ZGGA</td>
<td>0.007***(+4.826)</td>
<td>0.005***(+10.592)</td>
<td>0.96</td>
<td>2.22</td>
<td>333156</td>
</tr>
<tr>
<td>KHSW</td>
<td>-0.013***(-2.75)</td>
<td>-0.056***(-4.075)</td>
<td>0.38</td>
<td>2.18</td>
<td>KHSW</td>
<td>0.007***(+3.65)</td>
<td>0.006***(+12.67)</td>
<td>0.95</td>
<td>2.35</td>
<td>293061</td>
</tr>
<tr>
<td>LDGD</td>
<td>-0.146***(-5.88)</td>
<td>-0.143***(-2.92)</td>
<td>0.31</td>
<td>2.19</td>
<td>LDGD</td>
<td>0.009***(-2.79)</td>
<td>0.006***(-7.29)</td>
<td>0.92</td>
<td>2.21</td>
<td>100603</td>
</tr>
<tr>
<td>SZTY</td>
<td>-0.015(-1.07)</td>
<td>-0.137***(-3.507)</td>
<td>0.32</td>
<td>2.16</td>
<td>SZTY</td>
<td>-0.005(-1.70)</td>
<td>-0.001(-1.16)</td>
<td>0.87</td>
<td>2.21</td>
<td>253979</td>
</tr>
<tr>
<td>JRXG</td>
<td>-0.106*(-2.03)</td>
<td>-0.199***(-1.726)</td>
<td>0.20</td>
<td>2.16</td>
<td>JRXG</td>
<td>-0.001(-0.20)</td>
<td>0.006***(-4.03)</td>
<td>0.82</td>
<td>2.27</td>
<td>64553</td>
</tr>
<tr>
<td>FXXC</td>
<td>-0.581***(-5.98)</td>
<td>-0.658***(-23.18)</td>
<td>0.27</td>
<td>2.19</td>
<td>FXXC</td>
<td>-0.006(-1.32)</td>
<td>-0.003*(-2.10)</td>
<td>0.75</td>
<td>2.18</td>
<td>104802</td>
</tr>
<tr>
<td>RZG</td>
<td>-0.010***(-4.20)</td>
<td>-0.004(-4.26)</td>
<td>0.13</td>
<td>2.04</td>
<td>RZG</td>
<td>-0.004***(-2.40)</td>
<td>0.006(21.85)</td>
<td>0.99</td>
<td>2.13</td>
<td>140940</td>
</tr>
<tr>
<td>ZFSY</td>
<td>-0.030***(-5.52)</td>
<td>-0.025***(-18.34)</td>
<td>0.18</td>
<td>2.21</td>
<td>ZFSY</td>
<td>0.009***(-3.77)</td>
<td>0.004***(-6.80)</td>
<td>0.96</td>
<td>2.27</td>
<td>243496</td>
</tr>
<tr>
<td>BHGF</td>
<td>-0.047***(-4.61)</td>
<td>-0.075***(-30.51)</td>
<td>0.35</td>
<td>2.19</td>
<td>BHGF</td>
<td>0.001(0.02)</td>
<td>0.003***(-3.66)</td>
<td>0.93</td>
<td>2.23</td>
<td>224913</td>
</tr>
</tbody>
</table>

NOTE: The t-statistics are in the parentheses, ***, **, * means the coefficient is significant at 1%, 5%, 10% level separately.

REFERENCES