Estimation of Market Resiliency from High-Frequency Micex Shares Trading Data

Nikolay Andreev

National Research University Higher School of Economics, Moscow, Russia.
email: nick.my.mail@gmail.com

Abstract This article presents an engineering approach to estimating market resiliency based on analysis of the dynamics of a liquidity index. The method provides formal criteria for defining a "liquidity shock" on the market and can be used to obtain resiliency-related statistics for further research and estimation of this liquidity aspect. The developed algorithm uses the results of a spline approximation for observational data and allows a theoretical interpretation of the results. The method was applied to real data resulting in estimation of market resiliency for the given period.

Keywords: liquidity, portfolio liquidation, resiliency, transaction costs, bid-ask spread.

JEL classification: C65; G12.

Introduction

Market liquidity is a point of interest for many practical applications. This paper demonstrates an approach to estimating one of the characteristic of liquidity – market resiliency. This concept was introduced by Kyle (1985) along with two other concepts, tightness and depth, and defined as the rate at which prices recover from the uninformative shock. One of the main applications of this work's result is estimating the minimal time interval between consequent trades during portfolio liquidation, as described, for example, in Almgren & Criss (1999). The main optimal condition of the approach is minimizing transaction costs by dividing the portfolio volume into N parts and liquidating one part per trade. The problem is that each trade will lead to a price impact and large transaction costs, thus making it ineffective to participate in the market immediately after that. Estimating market resiliency will prove important in measuring the time of replenishment for the market and, therefore, the minimal interval between trades.

Measuring resiliency is a relatively new field of research in financial engineering. One of the first approaches in literature was the so-called γ coefficient, the time of a market's returning to "normal" state. "Returning to normal" in this framework means that the bid-ask spread takes on a pre-shock value. Such a concept doesn't take into consideration the fact that for an illiquid market, returning to the same values of spread and price may not happen, but move to the new "normal" stationary state.

Another approach was developed in Large (2007), based on using parametric impulse response functions for different kinds of events in the market. In that framework, returning to "normal" state means near-zero values of impulse functions. However, the author indicates that both the bid and the ask have less than 20% of replenishment after the large order.

The approach introduced here uses historical information about MICEX share trades. We use historical data to define shock states as a significant deviation from common behavior both in the nearest past and the nearest future. The statistics obtained are used to define the longest period of continuous shock condition, which is later used as an estimator of market resiliency.

The paper proceeds as follows: Section 2 describes the formal criterion for defining shock states of the market and analyses the results, Section 3 concludes.

Method for Detecting Shock States of a Liquidity Indicator

In this section we provide an engineering approach to estimating market resiliency using high-frequency shares trading data. The method is based on analysis of a liquidity index (phase variable). In this work we focus on the Xetra Liquidity Measure, closely related to average price impact costs, as the variable. This index aggregates the market impact information on the bid and ask side of the limit order book. It describes the performance loss due to liquidity costs that occur during simultaneous opening and closing a position of volume *V*. Construction of the index is quite simple and can be obtained from the following algorithm: for each moment *t* let $B_t(V)$ be the aggregate cost of opening a position of volume *V*, $C_t(V)$ – the aggregate cost of closing a position of the same volume. Then, by Xetra Liquidity Measure at the moment *t* we mean

$$\text{Xetra Liquidity Measure}_t(V) = \frac{B_t(V) - C_t(V)}{V}.$$

The associated volume *V* must be rather large to avoid negligible fluctuations in the dynamics. In the following research we take *V equal to* half of the average traded volume during trading period. For more information about Xetra Liquidity Measure index see Gomber & Schweickert (2002).

In this framework we define "liquidity shock" as the deviation of the phase variable, hereafter $Y(t)$, from its typical behavior. By shock length we mean the time necessary to return to the normal state. Fig.1 shows $Y(t)$ dynamics for "Lukoil" shares for 30 minutes in the middle of the trading period (10th January, 2006).

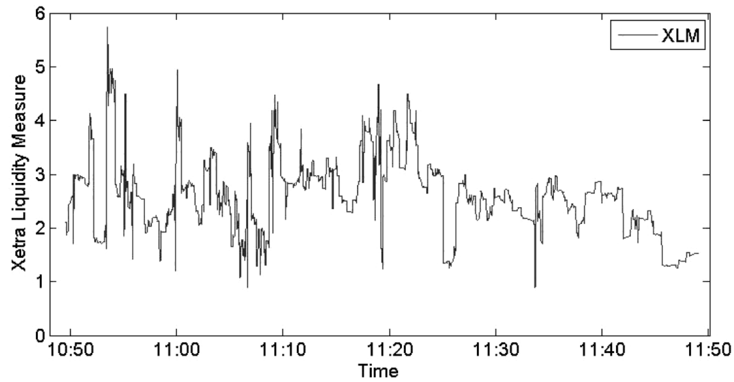

Fig. 1: Phase variable dynamics

This case already shows that intuition doesn't always allow one to detect shock states of the market (see, for example, peak at around 11:05 or 11:12). Thus a formal criterion is necessary to separate the normal and extraordinary behavior of the process. The remain of the chapter is divided into three parts

1. Estimating the trend;
2. Constructing a characteristic function for the given trajectory, and interpretation of the results;
3. Providing a criterion to detect irregular states in dynamics.

1. Estimation of common dynamics is necessary for further analysis because it allows one to neglect the influence of the global effects such as monotony or oscillation of the series. The results of the work hold under the following algorithm of defining trend $L(t)$:

Suppose we have observations of the underlying trajectory $(y_0, y_1,..., y_n) = (Y(t_0), Y(t_1),..., Y(t_n))$ at the discrete moments of time $t_0 < t_1 < ... < t_n \le T$. In this case $L(t)$ on [0, T] can be found as the solution of the following minimization problem:

$$\sum_{i=0}^{n} \alpha_i (L(t_i) - Y(t_i))^2 + \varepsilon \int_0^T (L''(s))^2 ds \to \inf_{L \in W},$$

where W is the so-called Sobolev-Hilbert space of functions with an absolutely continuous first derivative and second derivative from $L_2[0,T]$. The a priori parameter ε is positive and represents the tradeoff between fidelity and smoothness (a larger values mean smoother curves). Weights α_i are found as $\alpha_i = \frac{c}{1 + |X(t_i) - \overline{X}|}$, where c is a positive constant to secure the normalization condition $\alpha_0 + \alpha_1 + ... + \alpha_n = 1$. It is shown in Wahba (1990) that the solution of the problem is a piecewise-polynomial function. Figure 2 shows the solution $L(t)$ (dashed line) for sufficiently large ε.

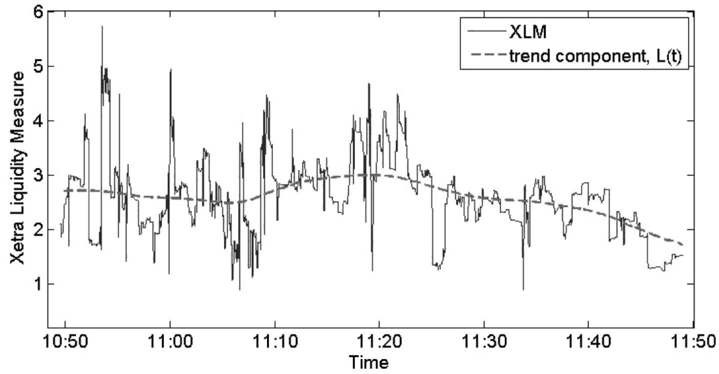

Fig. 2: Trend and trajectory of the phase variable

It is worth mentioning that the algorithm converges to the least-squares method as $\varepsilon \to +\infty$.

2. Constructing a characteristic function of the given series requires that some assumptions hold. We formally assume that $(y_0, y_1, ..., y_n) = (Y(t_0), Y(t_1), ..., Y(t_n))$ are the noised observations of a trajectory of some general stochastic process $F(t)$. The proposed model is

$$F(t) = L(t) + bX(t), \quad t \in [0, T],$$

$$Y(t_i) = F(t_i) + \eta_i, \quad i = 0, 1, ..., n,$$

where $L(t)$ is a stationary component found in the previous stage;

b is an unknown positive constant;

X (t) is the integrated Wiener process, i.e. Gaussian process with zero mean and known covariance function:

$$EX(t) = 0, \; R(t, s) = EX(t)X(s) = \int_0^T (t-u)_+ (s-u)_+ du,$$

where $x_+ = \max(x, 0)$;

$\eta_0, \eta_1, ..., \eta_n$ are i.i.d. random variables with normal distribution $N(0, \sigma^2)$. Under the assumptions the following statement holds:

Theorem (Kimeldorf & Wahba, 1970): let $\hat{F}(t)$ be the minimum variance, unbiased linear estimate of $F(t)$ given $(y_0, y_1, ..., y_n) = (Y(t_0), Y(t_1), ..., Y(t_n))$. Let $f_\varepsilon(t)$ be the solution of the minimization problem

$$\sum_{i=0}^{n} \alpha_i \left(f(t_i) - Y(t_i)\right)^2 + \varepsilon \int_0^T \left(f''(s)\right)^2 ds \to \inf_{f \in W}, \quad \varepsilon = \frac{\sigma^2}{b^2},$$

where W is the Sobolev-Hilbert space of functions with absolutely continuous first derivative and second derivative from $L_2[0,T]$. *Then* $\hat{F}(t) = f_\varepsilon(t)$.

From here on it is convenient to think of $f_\varepsilon(t)$ as a function of two arguments *t* and ε: $f_\varepsilon(t) \equiv f(t,\varepsilon)$. Using the statement it follows that, with the assumption of fixed $\varepsilon = \varepsilon_0$, the residual will be

$$\begin{aligned} &E\left(\hat{F}(t) - F(t) \,|\, \varepsilon = \varepsilon_0\right)^2 = \\ &\left(f(t,\varepsilon_0) - L(t)\right)^2 + 2b\left(f(t,\varepsilon_0) - L(t)\right)EX(t) + b^2 EX^2(t) = \\ &= const + g^2(t,\varepsilon_0)\ , \end{aligned}$$

where $g(t,\varepsilon) = f(t,\varepsilon) - L(t)$ is the deviation from the “mean” function.

However, real data does not allow one to directly use the results of the Theorem, due to the unknown parameters *b*, dispersion σ^2, and, therefore, the regularization parameter ε. This problem can be avoided by allowing only a priori information about ε but **not** its **exact value**. Assuming that we know some information about **the possible values** of parameter, it is convenient to use logical interpretation of probability and consider ε as a random variable with a priori distribution. In the case of no exogenous information available, the only property of the regularization parameter is positivity. Thus the most appropriate distribution is exponential with mean λ based on the fact that among distributions on positive semi axis and with fixed mean the exponential possesses the maximal entropy. Empirical studies show that the method is robust to the choice of λ which allows a rough estimation of the parameter according to the sufficiency of the results. In this demonstration $\lambda = 1$ was used.

The stochastic nature of ε leads to finding the **expected residual** $E_\varepsilon E\left(\hat{F}(t) - F(t)\right)^2$ of the estimation:

$$E_\varepsilon E\left(\hat{F}(t) - F(t)\right)^2 = const + \int_0^{+\infty} \lambda e^{-\lambda\varepsilon} g^2(t,\varepsilon) d\varepsilon = const + \psi(t)\ ,$$

$$\psi(t) = \lambda \int_0^{+\infty} e^{-\lambda\varepsilon} g^2(t,\varepsilon) d\varepsilon.$$

The obtained function $\psi(t)$ is non-negative and has sharp deviations when the expected residual is at its maximum. Therefore, at such moments, the estimation of the phase dynamics by observations is most difficult, i.e. the variable’s behavior aberrates from usual and predictable, interpreted in this framework as a shock state. Only the relative amplitude of $\psi(t)$ is important, so it is computationally easier to work with normalized values of the function. Figure 3 demonstrates the behavior of the original trajectory and the corresponding characteristic function.

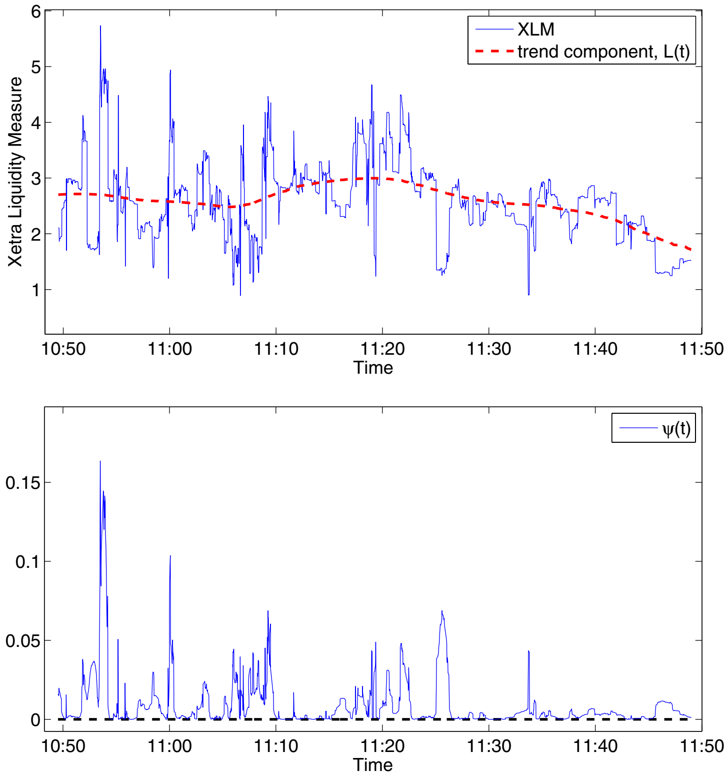

Fig. 3: Phase variable dynamics and characteristic function

The obtained results show that the stationary dynamics of the series correspond to near-zero values of $\psi(t)$. All "obvious" shock states match the function's deviations with high amplitude.

The method can be improved through classifying deviations by either ascending or descending behavior of the trajectory (hereafter upper and lower shocks correspondingly). In particular, a point of interest is detecting upward aberrations (lack of liquidity at the market), which is a direct consequence of the economic interpretation of the phase variable (Xetra Liquidity Measure). The final result of the resiliency's estimation will be based on this class of shocks.

The direction of shock for each moment t can be approximately established by using the sign of the deviation function $g(t,\varepsilon)$. For the case of the stochastic nature of ε we follow the same logic as before and derive the sign function

$$\chi(t) = sign\left\{\lambda \int_0^{+\infty} e^{-\lambda\varepsilon} g(t,\varepsilon) d\varepsilon\right\}.$$

Then $\chi(t)=1$ for upper shocks and $\chi(t)=-1$ otherwise. From now on the characteristic function of the trajectory can be written as

$$\psi(t) = \chi(t)\cdot\lambda \int_0^{+\infty} e^{-\lambda\varepsilon} g^2(t,\varepsilon) d\varepsilon.$$

It has the same properties as the previous one except for non-negativity. Aberrations of $\psi(t)$ in the positive half plane mean a decrease in liquidity. Figure 4 demonstrates the behavior of the original trajectory and the renewed characteristic function.

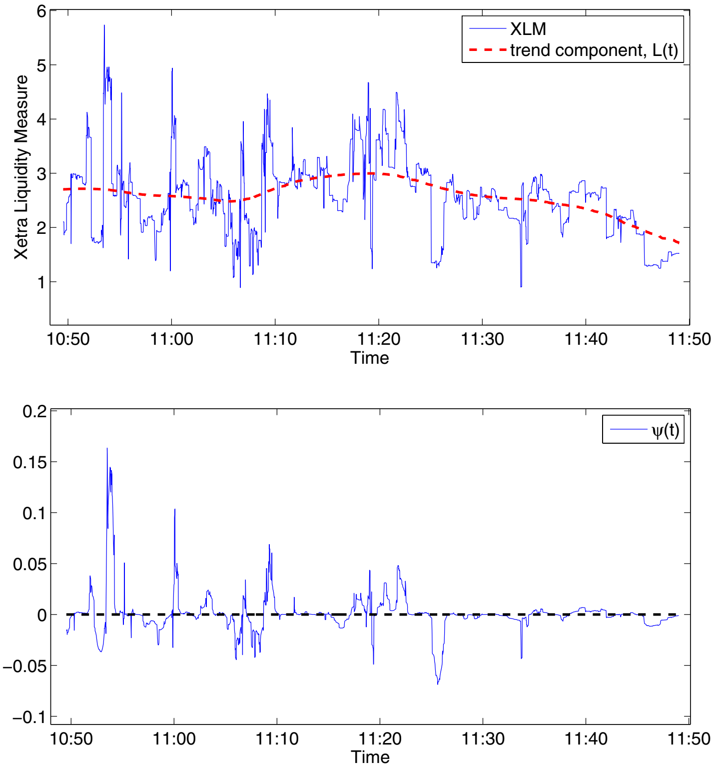

Fig. 4: Phase variable dynamics and characteristic function

$\psi(t)$ already allows visible detection of both types of shocks, and in particular a lack in liquidity. The next part of the section will provide an algorithm for an automatic strategy.

3. The formal criterion of shock will be based on constructing feasible bounds for the characteristic function. Overrunning these bounds will indicate shock behavior of the market. Instead of the continuous function $\psi(t)$ we consider a vector of its values $(\psi(t_0),\psi(t_1),...,\psi(t_n))$ for discrete moments of time $t_0 < t_1 < ... < t_n \le T$ (in this work the time-step is one second). The approach will be illustrated for the upper-shock bound but can be easily extrapolated for the other class.

For this purpose we consider only $(\psi_0,\psi_1,...,\psi_k) = (\psi(t_0'),\psi(t_1'),...,\psi(t_k'))$, where moments $t_0',t_1',...,t_k'$ are such that $\psi(t_i') \ge 0$. The upper-shock bound $m(t)$ can be constructed with various methods. The upper confidence level concept is proposed as rather simple and simultaneously efficient. We formally assume that

$$\psi(t_i) = l(t_i) + v_i,\;\; v_i - i.i.d.,\;\; N(0,\sigma_\psi^2),$$

which provides the following formula for $m(t)$:

$$m(t_i) = l(t_i) + q_\alpha \sigma_\psi,$$

where $l(t)$ can be defined with a spline approach for observations $(\psi_0,\psi_1,...,\psi_k)$;

σ_ψ^2 is the sample variance of the series $(\psi_0,\psi_1,...,\psi_k)$;

q_α is the fractile of the normal distribution $N(0,\sigma_\psi^2)$ for $\alpha\%$ level.

The criterion of the shock moment can be formally written as

$\{\tau$ is an upper-shock moment$\} \Leftrightarrow \psi(\tau) > m(\tau)$.

Comment: In many cases the aberrations of $\psi(t)$ have extremely high amplitude, thus leading to overestimation of the sample variance and not sensitive bounds. This problem can be avoided by conducting several preliminary iterations of the algorithm to remove high-amplitude moments t_i' from the associated set.

Fig.5 demonstrates the graphics of the characteristic function and the obtained bounds for a 99% confidence level. Fig.6 shows the trajectory of phase variable with marked shock states.

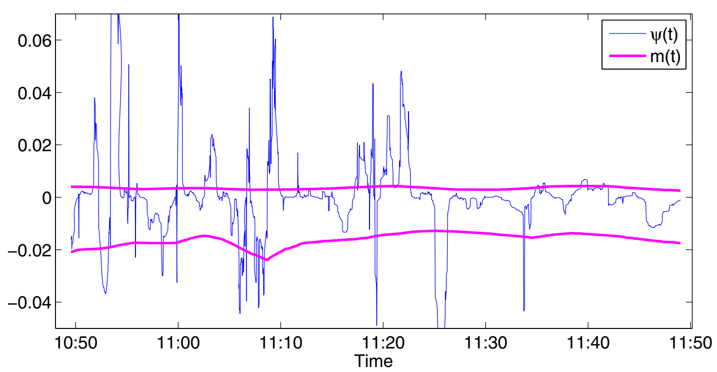

Fig. 5: Characteristic function and feasible bounds

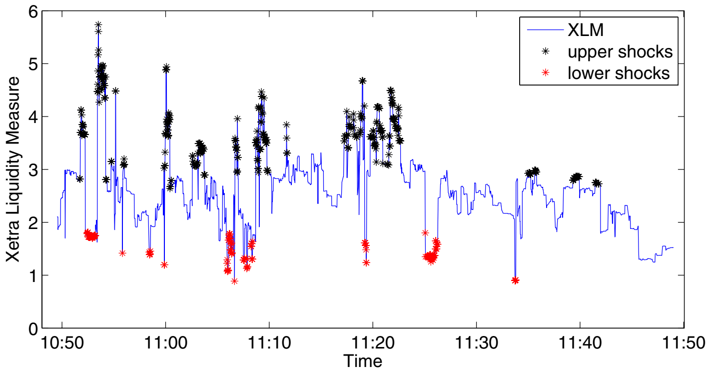

Fig. 6: Original trajectory with marked shocks

Market resiliency can now be estimated according to the statistics of continuous shock-periods. As for upper shocks, Table 1 shows that with 99.2% confidence, a 50 second period proves long enough for the market to recover after a shock. This estimate can be successfully used as a minimal time interval between consequent trades during piecewise liquidation strategy.

Table 1: Length of upper-shock states and percentage during 10th January, 2006, for "Lukoil" shares

Shock length	Percentage
< 50 seconds	99.2%
< 45 seconds	97.6%
< 30 seconds	88.1%
< 6 seconds	54.0%
< 5 seconds	49.2%

Conclusion

To quantify resiliency, a method for detecting shock states of the market was proposed. It allows automatic identification of aberrations in terms of a phase trajectory as a characteristic of liquidity. The algorithm is based on a smooth approximation approach and does not impound conditions on input data (long-term stable periods, sufficient period of time etc.). The robustness of the method and the easy interpretation of the results, correlating with the intuitive definition of shock, make it appropriate for obtaining statistics from historical data to estimate market resiliency. The method was tested on MICEX liquid shares trading data. For a period of one trading day it was shown that with a high (99.2%) level of confidence, 50 seconds are enough for the market to restore after an uninformative liquidity shock. Similar results can be derived for other periods and shares. But returning to a previous value of transaction costs, and thus liquidity level, is not a usual event at the market, which gives the proposed method an advantage in practical use.

References

Almgren, R., & Chriss, N.A. (1999). Optimal Execution of Portfolio Transactions. Retrieved from http://www.math.nyu.edu/faculty/chriss/optliq_f.pdf

Gomber, P., & Schweickert, U. (2002). The Market Impact - Liquidity Measure in Electronic Securities Trading, Working Paper. Retrieved from http://deutsche-berse.com/dbag/dispatch/en/binary/gdb_content_pool/imported_files/public_files/10_downloads/31_trading_member/10_Products_and_Functionalities/40_Xetra_Funds/30_Xetra_Liquidity_Measure/liq_wph.pdf

Kimeldorf, G.S., & Wahba, G. (1970). Spline Functions and Stochastic Processes. The Indian Journal of Statistics. Series A, Vol.32, No 2,173-180.

Kyle, A. (1985). Continuous Auctions and Insider Trading. Econometrica, 53, 1315-1336.

Large, J. (2007). Measuring the resiliency of an electronic limit order book. Journal of Financial Markets, 10, 1-25.

Wahba, G.(1990). Spline Models for Observational Data. Philadelphia, PA: SIAM.

http://www.springer.com/978-3-642-27930-0

Market Risk and Financial Markets Modeling
Sornette, D.; Ivliev, S.; Woodard, H. (Eds.)
2012, VIII, 268 p., Hardcover
ISBN: 978-3-642-27930-0