

Joanna OLBRYŚ<sup>1</sup>  
Michał MURSZTYN<sup>1</sup>

## MEASUREMENT OF STOCK MARKET LIQUIDITY SUPPORTED BY AN ALGORITHM INFERRING THE INITIATOR OF A TRADE

The aim of this study is to assess and analyse selected liquidity/illiquidity measures derived from high-frequency intraday data from the Warsaw Stock Exchange (WSE). As the side initiating a trade cannot be directly identified from a raw data set, firstly the Lee–Ready algorithm for inferring the initiator of a trade is employed to distinguish between so-called buyer- and seller-initiated trades. Intraday data for fifty-three WSE-listed companies divided into three size groups cover the period from January 3, 2005 to June 30, 2015. The paper provides an analysis of the robustness of the obtained results with respect to the whole sample and three consecutive subsamples, each of equal size: covering the pre-crisis, crisis, and post-crisis periods. The empirical results turn out to be robust to the choice of the period. Furthermore, hypotheses concerning the statistical significance of coefficients of correlation between the daily values of three liquidity proxies used in the study are tested.

**Keywords:** *liquidity, algorithm for inferring the initiator of a trade, intraday data*

### 1. Introduction

Classical finance theory is based on the assumption of a perfectly liquid market, where any security can be traded at no cost at any time, and agents take prices as given [1]. However, recently there has been a growing understanding of the crucial roles played by, e.g., liquidity, trading volume, bid/ask spread and other transaction costs. Bekaert et al. [2], among others, point out that liquidity/illiquidity is notably important for asset pricing. Illiquid assets and assets with high transaction costs are often traded at a low price relative to their expected cash flows. Therefore, the measurement of sys-

---

<sup>1</sup> Białystok University of Technology, Faculty of Computer Science, ul. Wiejska 45A, 15-351 Białystok, Poland, e-mail address: j.olbrys@pb.edu.pl

tematic risk should incorporate the costs of illiquidity (e.g., [25, 26]). Due to the importance of this problem, investors should recognize whether they have to take the risk of illiquidity into consideration in their financial decisions concerning the choice and diversification of portfolios.

The main goal of this paper is to assess and analyse selected liquidity/illiquidity measures based on intraday data for fifty-three WSE-listed companies divided into three size groups. Measuring liquidity on the WSE is an important and problematic subject. For example, Nowak and Olbryś [24] documented cross-time and cross-security patterns in non-trading among WSE-traded stocks. Their empirical results reveal that a large number of companies exhibit the phenomenon of substantial non-trading, which means a lack of transactions over a particular period when the WSE is open for trading.

The high-frequency intraday data rounded to the nearest second cover the period from January 3, 2005 to June 30, 2015. As the initiator of a trade cannot be directly identified from the raw data set, firstly the Lee–Ready [20] algorithm for classifying the initiator of a trade is employed to distinguish between so-called buyer- and seller-initiated trades [30]. Moreover, the paper provides an analysis of the robustness of the obtained results with respect to the whole sample and three consecutive sub-samples of equal size: covering the pre-crisis, crisis, and post-crisis periods. The Global Financial Crisis (GFC) on the WSE is a formally defined set based on the papers [28, 29], in which the Pagan and Sossounov [32] method for the formal statistical identification of market states was employed.

To the best of the authors' knowledge, the empirical results on the WSE presented here are novel and have not been reported in the literature thus far. The remainder of the study is organized as follows: Section 1 presents the Lee–Ready [20] rule for inferring the initiator of a trade. Section 2 describes the methodological background concerning the measurement of liquidity/illiquidity using intraday data. Section 3 presents and discusses the empirical results for the data from the WSE. The last section summarizes the main findings, together with a conclusion.

## **2. Algorithms for inferring the initiator of a trade**

High frequency financial data are important in studying a variety of issues related to trading processes and the microstructure of markets. To calculate various liquidity/illiquidity measures using intraday data, it is essential to recognize the side initiating the transaction and to distinguish between so-called buyer- and seller-initiated trades. The WSE is classified as an order-driven market with an electronic order book, but information regarding the best bid and ask price is not publicly available. In fact, even the non-proprietary financial databases that provide information on trades and quotes do not identify the initiator of a trade. As a consequence, researchers rely on indirect classification

rules to infer the initiator of a trade. Various classification procedures of this type are described in the literature, but the Lee–Ready [20] algorithm (LR) remains the most frequently used [4, p. 468]<sup>2</sup>.

Table 1. The Lee–Ready (LR) algorithm for inferring the initiator of a trade

I stage	
Trade is classified as buyer-initiated if $P_t > P_t^{\text{mid}}$	Trade is classified as seller-initiated if $P_t < P_t^{\text{mid}}$
If $P_t = P_t^{\text{mid}}$ then:	
II stage	
Trade is classified as buyer-initiated if $P_t^{\text{mid}} > P_{t-1}$	Trade is classified as seller-initiated if $P_t^{\text{mid}} < P_{t-1}$
When $P_t^{\text{mid}} = P_{t-1}$ , the decision is taken according to the sign of the last non-zero price change. If $P_t > P_{t-k}$ then trade is classified as buyer-initiated, if $P_t < P_{t-k}$ then it is classified as seller-initiated.	

Source: [30, p. 42].

Table 1 presents details concerning the LR procedure. The midpoint price  $P_t^{\text{mid}}$  at time  $t$  is calculated as the arithmetic mean of the best ask price  $P_t(a)$  and the best bid price  $P_t(b)$  at time  $t$ :  $P_t^{\text{mid}} = \frac{P_t(a) + P_t(b)}{2}$ . Considering that the bid and ask prices are not made public on the WSE, the midpoint price  $P_t^{\text{mid}}$  at time  $t$  is approximated by the arithmetic mean of the lowest price  $P_t^L$  and the highest price  $P_t^H$  at time  $t$ , which approximate the best ask price and the best bid price, respectively. The transaction price  $P_t$  at time  $t$  is approximated by the closing price. The opening trade is treated as being unclassified according to the LR procedure.

In this paper, the LR method is employed, as Olbryś and Mursztyn [30] indicated that the LR algorithm performs quite well for data from the WSE. The empirical results turn out to be robust to the choice of the sample and do not depend on a firm's size<sup>3</sup>. Table 2 presents the average percentage values of classified and unclassified trades for the 53 companies considered as a whole and the three size groups (large, medium-sized, and small companies), for the whole sample period and three consecutive subsamples, each of equal size<sup>4</sup>. The empirical findings indicate that the percentage of unclassified trades is rather low, regardless of firm size and the choice of the period, which is consistent with the literature.

<sup>2</sup>For a brief literature review concerning various trade classification rules, see, e.g., [30, p. 39–42].

<sup>3</sup>For details concerning the C++ program for the LR classification of trades, see [30, p. 48].

<sup>4</sup>Details concerning the companies and data used in this study are described in Section 4.

Table 2. Average percentage values of classified and unclassified trades for the large, medium, and small groups (the Lee–Ready procedure)

Period	Group	Total number of records	Percentage of trades		
			Buyer-initiated	Seller-initiated	Unclassified
Whole sample	all	22 817 300	48.37	45.80	5.83
	large	19 828 145	48.88	46.91	4.21
	medium	2 359 773	47.38	44.61	8.01
	small	629 382	48.92	44.70	6.38
Pre-crisis	all	3 284 945	49.05	45.50	5.45
	large	2 311 742	48.61	46.17	5.22
	medium	683 180	48.65	44.30	7.05
	small	290 023	51.39	45.95	2.66
Crisis	all	3 716 098	46.90	46.84	6.26
	large	3 110 255	47.32	47.57	5.11
	medium	471 879	45.77	46.35	7.88
	small	133 964	48.01	45.50	6.49
Post-crisis	all	4 191 750	47.75	44.37	7.88
	large	3 664 509	48.18	45.15	6.67
	medium	432 739	47.47	43.78	8.75
	small	94 502	46.94	43.06	10

All – 53 companies, large – 27 companies, medium-sized – 18 companies, small – 8 companies. Source: Authors' calculations.

### 3. Some liquidity proxies derived from intraday data

Direct measurement of, e.g., liquidity, bid/ask spreads or other trading costs is difficult or even impossible as intraday transaction data are not available free of charge in the case of most emerging stock markets (e.g., [2, 21, 23, 25, 26]). The literature presents many alternative measures of stock market liquidity/illiquidity based on intraday transaction data, as well as indicators of imbalance in market orders (e.g., [6–8, 13, 17, 22, 23, 25, 33, 35, 38]).

Three alternative estimates of liquidity/illiquidity derived from intraday data are employed: (1) the percentage order ratio as an indicator of order imbalance, (2) percentage realized spread, and (3) percentage proxy of price impact. To calculate these measures, it is essential to recognize the side that initiates a transaction and to distinguish between buyer- and seller-initiated trades by using an algorithm to infer the initiator of trade in the first step of analysis. Moreover, both the realized spread and price impact proxies are treated as components of the effective spread, and they are calculated over a time interval that begins at the moment of a buyer- or seller-initiated transaction.

For example, Goyenko et al. [13] employ a five minute interval and the subscript  $t + 5$  indicates trade five minutes after trade at time  $t$ . Chakrabarty et al. [3] use the subscript  $t + 10$  which indicates trade ten minutes after trade at time  $t$ . Theissen [36] proposes a more general approach and the subscript  $t + \tau$ . In this study, the subscript  $t + 5$  indicates the fifth trade after the  $t$ -th trade (made at moment  $t$ ), as a large number of the WSE-listed companies exhibit a substantial degree of non-trading, i.e., there is a lack of transactions over a particular period when the WSE is open for trading [24].

### 3.1. Indicator of order imbalance

Order imbalance has a significant influence on stock liquidity, considerably more important even than volume. Therefore, indicators of order imbalance could be employed among other measures of liquidity and trading activity to estimate liquidity. The literature proposes various proxies for order imbalance (e.g., [5, 7, 8, 18, 23, 27, 31, 33, 38]). The percentage order ratio (% OR) is employed as an indicator of imbalance in daily orders:

$$\% \text{ OR} = \frac{\left| \sum_{i=1}^m V\text{Buy}_i - \sum_{j=1}^k V\text{Sell}_j \right|}{\sum_{n=1}^N V_n} \times 100 \quad (1)$$

where the sums  $\sum_{i=1}^m V\text{Buy}_i$ ,  $\sum_{j=1}^k V\text{Sell}_j$ ,  $\sum_{n=1}^N V_n$  denote the daily cumulative volume of trading related to transactions classified as buyer- or seller-initiated trades, and daily cumulative volume of trading for all transactions, respectively. The OR indicator (1) captures imbalance in the market, since it rises as the difference in the numerator grows. A high value of the order ratio denotes low liquidity. Conversely, a small value of the order ratio denotes high liquidity. The OR indicator is equal to zero when the numerator is equal to zero. This happens when the daily cumulative volumes of trading related to transactions classified as buyer- and seller-initiated trades, respectively, are equal. Moreover, the value of the daily order ratio is defined to be equal to zero in the following two cases: (1) when all of the transactions within a day are unclassified, or (2) when the total volume of daily trading, the denominator, is equal to zero.

### 3.2. Realized spread

The realized spread is a temporary component of the effective spread, which is defined as the amount earned by a dealer or other immediate supplier (e.g., [15, 36]). The

realized spread is sometimes referred to as the component of price reversal, since a dealer makes a profit only if the price reverses. The percentage value of the realized spread (% RealS) is given by Eq. (2):

$$\% \text{ RealS}_t = \begin{cases} 200 \ln \frac{P_t}{P_{t+5}}, & \text{when trade } t \text{ is classified as buyer-initiated} \\ 200 \ln \frac{P_{t+5}}{P_t}, & \text{when trade } t \text{ is classified as seller-initiated} \end{cases} \quad (2)$$

where the transaction price  $P_t$  at moment  $t$  is approximated by the closing price. The price  $P_{t+5}$  is the closing price of the fifth trade after trade  $t$ . % RealS at moment  $t$  is equal to zero when  $P_t = P_{t+5}$ . The post-trade revenues earned by a dealer (or any other supplier of liquidity) are estimated on the basis of actual post-trade prices. The value of the daily percentage realized spread is calculated as a volume-weighted average of the percentage realized spreads computed over all the trades within a day. The value of the daily percentage realized spread is defined to be equal to zero when all of the transactions within a day are unclassified.

### 3.3. A proxy for price impact

According to the literature, a proxy for price impact measures the sensitivity of a stock's price to its trades [35, p. 1495], and most researchers derive price impact from intraday transaction data (e.g., [3, 9, 38]). Kyle [19] provides a theoretical model for such a measure based on the adverse information conveyed by a trade. Price impact could be defined as the increase (decrease) in the quote midpoint over a time interval beginning at the time of a buyer- (seller-) initiated trade. This is a permanent price change for a given transaction, or equivalently, a permanent component of the effective spread (e.g., [13, p. 156]).

The percentage value of price impact (% PI) focuses on the change in the quote midpoint after a signed trade and is given by Eq. (3):

$$\% \text{ PI}_t = \begin{cases} 200 \ln \frac{P_{t+5}^{\text{mid}}}{P_t^{\text{mid}}} & \text{when trade } t \text{ is classified as buyer-initiated} \\ 200 \ln \frac{P_t^{\text{mid}}}{P_{t+5}^{\text{mid}}} & \text{when trade } t \text{ is classified as seller-initiated} \end{cases} \quad (3)$$

where  $P_t^{\text{mid}}$  is the midpoint price at moment  $t$ , while  $P_{t+5}^{\text{mid}}$  is the quote midpoint five trades after trade  $t$ . The % PI at moment  $t$  is equal to zero when  $P_t^{\text{mid}} = P_{t+5}^{\text{mid}}$ . The proxy for daily percentage price impact is calculated as a volume-weighted average of the estimates of percentage price impact computed over all the trades within a day. The value of the daily percentage price impact is defined to be equal to zero when all of the transactions within a day are unclassified.

#### 4. Description of the data and empirical results for the Warsaw Stock Exchange

We utilize a database containing high-frequency data rounded to the nearest second (available at [www.bossa.pl](http://www.bossa.pl)) for fifty-three WSE-listed stocks divided into three size groups, for the period from January 3, 2005 to June 30, 2015. When forming the data base, we included only those securities which had existed on the WSE for the whole sample period from December 31, 2004, and had not been suspended. All of the companies contained in this database (147) were sorted according to their market capitalization at the end of each year. Next, the stocks were divided into three size groups based on the following categorisation: the bottom 30% (small companies), the middle 40% (medium-sized companies), and the top 30% (large companies) [10]. Companies that remained in the same group for the whole of the period investigated were selected. In this way, 53 WSE companies were classified into three separate groups, specifically: 27 firms into the large group, 18 firms into the medium group, and 8 firms into the small group [24].

The dataset is large and contains the opening, high, low and closing (OHLC) prices, and volume for a security over one unit of time. For example, considering just trading days, during the whole sample period, there are 3 959 406 records for the most liquid Polish company, the KGH dataset. Therefore, special programs in the C++ programming language have been implemented to reduce the time required for calculations.

To verify the robustness of the empirical results, analysis was applied to the whole sample (2626 trading days) and three consecutive periods each of equal length (436 trading days): (1) the pre-crisis period, September 6, 2005 to May 31, 2007, (2) the crisis period, June 1, 2007 to February 27, 2009, and (3) the post-crisis period March 2, 2009 to November 19, 2010 [31]. The Global Financial Crisis on the WSE is a formally defined dataset based on the papers [28, 29], in which the Pagan and Sossounov [32] procedure for the formal statistical identification of market states was employed. Precise detection of market states is crucial, due to many practical implications. Among other things, the issue concerning the existence of interaction between stock market declines and market liquidity is very important (e.g., [14]).

#### 4.1. Empirical results for the order ratio on the WSE

The percentage order ratio (1) was utilized as an indicator of order imbalance. In the first step, we calculated the daily cumulative volume of trade related to transactions classified as buyer- and seller-initiated trades individually, as well as daily cumulative volume of trade for all transactions (including those unclassified), for each WSE-listed company with respect to its size group (i.e., large, medium, or small, as appropriate). In the second step, the average value of the daily percentage order ratio was approximated. The empirical results are reported in Table 3.

Table 3. The average daily percentage order ratio (% OR)

L	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>	M	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>	S	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>
BHW	38.4	46.9	47.0	49.2	ALM	42.9	38.7	46.6	36.8	APL	29.0	31.4	32.5	29.8
BPH	40.1	32.9	40.7	40.0	AMC	37.3	37.3	40.0	27.0	BDL	29.2	25.9	24.6	29.5
BNP	31.0	26.7	38.3	16.0	ATG	43.7	42.8	47.3	50.3	EFK	42.3	35.9	42.4	47.9
BOS	34.2	30.9	27.9	34.7	ATM	44.3	45.2	43.3	42.2	ENP	37.8	29.8	31.5	37.5
BDX	42.8	52.8	47.1	44.7	CNG	44.4	37.3	49.9	46.8	KMP	33.2	32.7	34.5	35.5
BZW	30.7	31.9	24.7	26.4	COL	37.2	47.0	40.1	24.8	MZA	35.9	32.8	39.0	33.1
DBC	43.6	41.3	49.0	41.3	IND	43.7	45.1	46.4	46.2	PLA	35.8	31.5	32.1	34.8
ECH	44.6	47.5	39.6	43.2	IPL	44.7	37.9	42.0	41.6	SME	41.5	37.6	39.1	43.2
GTN	27.4	25.9	29.1	25.2	LTX	33.7	28.5	28.4	33.0	Mean	35.6	32.2	34.5	36.4
GTC	30.1	33.1	24.8	26.0	MCI	24.9	24.4	24.7	17.4					
ING	48.0	57.6	53.2	43.4	MNI	34.1	26.8	27.8	40.8					
KTY	46.0	44.2	50.6	48.9	PEK	43.4	40.7	45.8	47.5					
KGH	17.0	16.7	18.9	18.5	PUE	41.5	42.1	38.6	42.6					
LPP	45.8	53.0	48.8	51.7	SKA	43.9	43.6	43.3	45.6					
MBK	29.3	39.6	28.0	24.2	STF	40.9	28.1	39.4	42.2					
MIL	35.2	38.0	39.0	29.8	STX	30.5	24.0	18.4	28.2					
MOL	46.6	43.6	49.2	49.5	TIM	43.4	38.4	46.2	47.5					
NET	36.5	29.2	42.3	39.4	VST	36.1	47.9	49.9	23.4					
OPL	21.4	19.6	20.1	21.7	Mean	39.5	37.5	39.9	38.0					
ORB	49.8	45.2	48.9	51.2										
PEO	21.1	24.2	21.0	21.1										
PKN	18.8	18.9	19.5	20.3										
PKO	19.5	23.3	20.5	19.7										
STP	44.7	43.0	47.5	45.8										
SNS	31.7	40.9	37.8	34.4										
TVN	27.2	28.4	25.2	26.1										
ZWC	38.8	41.1	41.6	42.3										
Mean	34.8	36.2	36.3	34.6										

This table is based on: (1) the whole sample period P<sub>1</sub>, (2) the pre-crisis period P<sub>2</sub>, (3) the Global Financial Crisis period P<sub>3</sub>, and (4) the post-crisis period P<sub>4</sub>. L – large, M – medium, S – small. Source: authors' calculations.





Table 4. The average daily percentage realized spread (% RealS)

L	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>
PKN	0.02	0.02	0.03	0.02
PKO	0.03	0.04	0.03	0.03
STP	0.14	0.16	0.22	0.16
SNS	0.12	0.12	0.20	0.20
TVN	0.08	0.09	0.08	0.08
ZWC	0.03	0.03	0.04	0.03
Mean	0.08	0.09	0.10	0.10

For explanation, see Table 3. Source: authors' calculations.

The average daily estimates of percentage realized spread are positive for almost all of the stocks from the three size groups, except for a few isolated cases. These findings are rather consistent with the literature because the existence of a bid/ask spread has several consequences for the properties of time series, and one of them is the bid/ask bounce (see, e.g., [34, 37]). According to Definition (2), the realized spread is, in fact, a percentage logarithmic rate of return. As a price reversal component of the bid/ask spread, the value of the realized spread is usually positive, since an investor makes a profit only if prices reverse. Therefore, a small absolute value of the realized spread indicates high liquidity, while a high absolute value of the realized spread denotes low liquidity. Moreover, one can observe that the results in Table 3 rather do not depend on a firm's size and turn out to be robust to the choice of the period.

### 4.3. Empirical results for the proxy for price impact on the WSE

The percentage price impact (3) was utilized as a proxy for a permanent price change for a given transaction. In the first step, we calculated the value of the % PI indicator (3) for each transaction classified as a buyer- or seller-initiated trade, for each WSE-listed company with respect to its size group. In the second step, the value of the daily percentage price impact was calculated as a volume-weighted average of the price impact estimates computed over all the classified trades within a day, for each company. Next, the average value of the proxy for daily percentage price impact was approximated. The empirical results are presented in Table 5.

The evidence reveals that the average daily estimates of price impact are negative in most cases, which is a probable consequence of the fact that both the realized spread and price impact proxies are treated as effectively components of the bid/ask spread which complement each other (see, e.g., [12, 15, 16]). We observe negative values of % PI close to zero for large companies with high liquidity and the largest market capitalization (namely KGH, PEO, PKN, PKO), regardless of the choice of subsample. However, the results reported in Table 5 rather do not depend on a firm's size in general.

Table 5. The average daily percentage price impact (% PI)

L	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>	M	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>	S	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>
BHW	-0.06	-0.11	-0.07	-0.07	ALM	-0.13	-0.14	-0.12	-0.09	APL	-0.07	-0.11	-0.10	-0.14
BPH	-0.07	-0.05	-0.11	-0.11	AMC	-0.08	-0.09	-0.14	-0.06	BDL	-0.15	-0.06	-0.12	-0.14
BNP	-0.01	0.03	-0.02	0.001	ATG	-0.04	0.02	0.04	0.01	EFK	-0.07	-0.12	-0.13	0.02
BOS	-0.01	-0.001	-0.01	-0.03	ATM	-0.10	-0.10	-0.16	-0.09	ENP	-0.13	-0.23	-0.12	-0.17
BDX	-0.06	0.04	-0.09	-0.05	CNG	-0.08	-0.11	0.004	-0.11	KMP	-0.14	-0.13	-0.19	-0.25
BZW	-0.04	-0.06	-0.05	-0.04	COL	-0.07	0.02	-0.08	-0.08	MZA	-0.12	-0.21	-0.15	-0.05
DBC	-0.04	-0.05	-0.03	-0.07	IND	-0.05	-0.04	-0.06	-0.14	PLA	-0.06	-0.11	-0.04	-0.14
ECH	-0.10	-0.04	-0.10	-0.14	IPL	-0.04	-0.06	-0.02	-0.03	SME	-0.06	-0.22	0.03	-0.06
GTN	-0.06	-0.11	-0.04	-0.04	LTX	-0.06	-0.03	-0.06	-0.08	Mean	-0.10	-0.17	-0.09	-0.13
GTC	-0.05	-0.07	-0.02	-0.05	MCI	-0.05	-0.08	-0.03	-0.01					
ING	-0.05	-0.08	-0.03	-0.04	MNI	-0.09	-0.10	-0.09	-0.06					
KTY	-0.08	-0.18	-0.09	-0.04	PEK	-0.08	-0.14	-0.16	-0.03					
KGH	0.000	0.000	-0.01	-0.007	PUE	-0.01	0.02	-0.02	-0.03					
LPP	-0.05	-0.01	-0.12	-0.08	SKA	0.000	-0.01	-0.07	-0.04					
MBK	-0.04	-0.10	-0.03	-0.04	STF	-0.06	-0.08	-0.07	-0.09					
MIL	-0.07	-0.09	-0.09	-0.05	STX	-0.10	-0.06	-0.03	-0.11					
MOL	-0.04	-0.11	-0.11	-0.02	TIM	-0.05	-0.10	-0.09	-0.01					
NET	-0.08	-0.11	-0.07	-0.08	VST	-0.09	-0.01	-0.12	-0.07					
OPL	-0.02	-0.03	-0.02	-0.04	Mean	-0.06	-0.06	-0.07	-0.06					
ORB	-0.08	-0.12	-0.07	-0.10										
PEO	-0.008	-0.02	-0.02	-0.02										
PKN	0.001	-0.004	-0.01	-0.004										
PKO	-0.01	-0.03	-0.14	-0.08										
STP	-0.07	-0.08	-0.10	-0.31										
SNS	-0.07	-0.03	-0.13	-0.14										
TVN	-0.05	-0.06	-0.06	-0.05										
ZWC	-0.01	-0.01	0.003	0.001										
Mean	-0.05	-0.06	-0.06	-0.07										

For explanation, see Table 3. Source: authors' calculations.

#### 4.4. Correlation analysis

In order to carry out a preliminary study of the interaction between these three proxies for liquidity, hypotheses concerning the statistical significance of correlation coefficients are tested. The basic idea is to apply Fisher's [11]  $z$ -transformation of a sample correlation coefficient to avoid the problem of a time series distribution being non-normal. Therefore, the OR/Reals and OR/PI correlations are represented by Fisher's  $z$ -transformation of the corresponding sample correlation coefficients. However, the values representing the Reals/PI correlations are the raw values of Pearson's sample correlation coefficient, because all of them are strongly associated with each other and Fisher's transformation is not necessary in such cases.

Table 6 reports the coefficients of correlation between the values of the daily percentage order ratio, daily percentage realized spread, and daily percentage price impact for the study group of fifty-three WSE-traded companies over the whole sample period ( $P_1$ ).

Table 6. Coefficients of correlation between the values of the daily percentage order ratio, daily percentage realized spread, and daily percentage price impact for 53 WSE-listed companies over the whole sample period from January 3, 2005 to June 30, 2015

L	OR /RealS	OR/PI	RealS /PI	M	OR /RealS	OR/PI	RealS /PI	S	OR /RealS	OR/PI	RealS /PI
BHW	0.034	-0.035	<i>-0.985</i>	ALM	<i>-0.056</i>	0.034	<i>-0.967</i>	APL	0.003	0.037	<i>-0.957</i>
BPH	<i>-0.016</i>	0.011	<i>-0.968</i>	AMC	<i>-0.067</i>	0.032	<i>-0.969</i>	BDL	0.009	0.003	<i>-0.960</i>
BNP	<i>-0.034</i>	<i>0.055</i>	<i>-0.982</i>	ATG	<i>-0.093</i>	<i>0.085</i>	<i>-0.979</i>	EFK	<i>-0.042</i>	0.038	<i>-0.967</i>
BOS	<i>-0.019</i>	0.037	<i>-0.968</i>	ATM	<i>-0.051</i>	<i>0.047</i>	<i>-0.984</i>	ENP	<i>-0.113</i>	<i>0.086</i>	<i>-0.959</i>
BDX	<i>-0.073</i>	<i>0.067</i>	<i>-0.972</i>	CNG	<i>-0.080</i>	<i>0.072</i>	<i>-0.985</i>	KMP	-0.006	0.013	<i>-0.970</i>
BZW	<i>0.041</i>	-0.038	<i>-0.979</i>	COL	<i>-0.158</i>	<i>0.135</i>	<i>-0.971</i>	MZA	-0.001	0.003	<i>-0.979</i>
DBC	<i>-0.079</i>	<i>0.071</i>	<i>-0.973</i>	IND	<i>-0.047</i>	<i>0.047</i>	<i>-0.981</i>	PLA	0.000	-0.004	<i>-0.960</i>
ECH	<i>0.026</i>	0.014	<i>-0.981</i>	IPL	<i>-0.076</i>	<i>0.076</i>	<i>-0.959</i>	SME	<i>-0.072</i>	<i>0.070</i>	<i>-0.965</i>
GTN	-0.033	0.002	<i>-0.959</i>	LTX	<i>-0.090</i>	<i>0.068</i>	<i>-0.962</i>	Median	-0.003	0.025	<i>-0.963</i>
GTC	-0.030	-0.026	<i>-0.978</i>	MCI	<i>-0.075</i>	<i>0.039</i>	<i>-0.950</i>				
ING	<i>-0.040</i>	0.028	<i>-0.981</i>	MNI	<i>-0.075</i>	<i>0.056</i>	<i>-0.951</i>				
KTY	-0.019	0.010	<i>-0.984</i>	PEK	<i>-0.068</i>	<i>0.117</i>	<i>-0.985</i>				
KGH	<i>0.133</i>	<i>-0.151</i>	<i>-0.942</i>	PUE	<i>-0.044</i>	<i>0.051</i>	<i>-0.978</i>				
LPP	-0.038	<i>0.041</i>	<i>-0.983</i>	SKA	<i>-0.035</i>	<i>0.218</i>	<i>-0.963</i>				
MBK	<i>0.071</i>	<i>-0.079</i>	<i>-0.985</i>	STF	<i>-0.035</i>	<i>0.009</i>	<i>-0.967</i>				
MIL	0.038	<i>-0.052</i>	<i>-0.973</i>	STX	<i>-0.058</i>	<i>0.026</i>	<i>-0.944</i>				
MOL	-0.020	0.026	<i>-0.990</i>	TIM	<i>-0.050</i>	<i>0.043</i>	<i>-0.979</i>				
NET	0.010	-0.022	<i>-0.977</i>	VST	<i>-0.104</i>	<i>0.071</i>	<i>-0.963</i>				
OPL	<i>0.057</i>	<i>-0.078</i>	<i>-0.958</i>	Median	<i>-0.067</i>	<i>0.053</i>	<i>-0.968</i>				
ORB	-0.024	0.002	<i>-0.990</i>								
PEO	<i>0.071</i>	<i>-0.078</i>	<i>-0.975</i>								
PKN	-0.032	0.025	<i>-0.957</i>								
PKO	<i>0.058</i>	<i>-0.091</i>	<i>-0.972</i>								
STP	<i>-0.055</i>	<i>0.047</i>	<i>-0.978</i>								
SNS	-0.007	0.014	<i>-0.967</i>								
TVN	<i>0.046</i>	<i>-0.061</i>	<i>-0.976</i>								
ZWC	<i>-0.083</i>	<i>0.087</i>	<i>-0.937</i>								
Median	-0.019	0.010	<i>-0.976</i>								

This table is based on the whole sample period  $P_1$ . The OR/RealS and OR/PI correlations are represented by Fisher's  $z$ -transform of correlation coefficients, while the RealS/PI correlations are represented by Pearson's correlation coefficient. The critical value for this correlation coefficient is equal to 0.038 at the 5% significance level (2626 daily observations). The significant correlation coefficients are marked in italics, Source: authors' calculations.

Tables 7 and 8 present the coefficients of correlation between the values of the daily percentage order ratio, daily percentage realized spread, and daily percentage price impact for the study group of fifty-three WSE-traded companies during the pre-crisis ( $P_2$ ) and crisis ( $P_3$ ) periods. Due to restrictions on space, the table based on the post-crisis period ( $P_4$ ), from March 2, 2009 to November 19, 2010 is not reported in the paper but is available upon request. However, the empirical results obtained for the post-crisis period are very similar to those presented in the tables.

Table 7. Coefficients of correlation between the values of the daily percentage order ratio, daily percentage realized spread, and daily percentage price impact for 53 WSE-listed companies in the pre-crisis period from September 6, 2005 to May 31, 2007

L	OR/RealS	OR/PI	RealS/PI	M	OR/RealS	OR/PI	RealS/PI	S	OR/RealS	OR/PI	RealS/PI
BHW	0.073	-0.085	-0.990	ALM	-0.044	0.053	-0.967	APL	-0.141	0.143	-0.957
BPH	-0.100	0.107	-0.973	AMC	-0.103	0.108	-0.958	BDL	0.100	0.006	-0.926
BNP	-0.087	0.137	-0.973	ATG	-0.100	0.109	-0.990	EFK	-0.247	0.260	-0.961
BOS	-0.061	0.080	-0.982	ATM	0.001	-0.011	-0.968	ENP	-0.022	0.025	-0.950
BDX	-0.062	0.074	-0.962	CNG	-0.170	0.143	-0.980	KMP	-0.022	-0.003	-0.951
BZW	0.040	-0.032	-0.986	COL	-0.253	0.259	-0.982	MZA	-0.113	0.137	-0.962
DBC	-0.128	0.118	-0.974	IND	-0.081	0.073	-0.971	PLA	0.105	-0.022	-0.967
ECH	-0.094	0.123	-0.967	IPL	-0.076	0.127	-0.846	SME	-0.059	0.055	-0.954
GTN	-0.014	-0.015	-0.973	LTX	-0.056	0.052	-0.938	Median	-0.040	0.040	-0.955
GTC	0.077	-0.067	-0.986	MCI	-0.020	0.043	-0.963				
ING	0.011	0.000	-0.976	MNI	-0.026	0.043	-0.931				
KTY	-0.021	0.024	-0.985	PEK	-0.119	0.123	-0.981				
KGH	0.151	-0.159	-0.977	PUE	-0.114	0.137	-0.981				
LPP	-0.056	0.067	-0.993	SKA	-0.072	0.095	-0.969				
MBK	0.113	-0.120	-0.989	STF	-0.018	0.017	-0.966				
MIL	0.001	0.027	-0.959	STX	0.061	-0.050	-0.952				
MOL	-0.015	0.027	-0.985	TIM	-0.118	0.098	-0.980				
NET	0.121	-0.124	-0.976	VST	-0.184	0.174	-0.984				
OPL	0.116	-0.114	-0.982	Median	-0.078	0.097	-0.968				
ORB	-0.071	0.074	-0.988								
PEO	0.096	-0.065	-0.975								
PKN	0.065	-0.065	-0.973								
PKO	0.048	-0.054	-0.983								
STP	-0.046	0.046	-0.985								
SNS	-0.187	0.179	-0.972								
TVN	-0.003	0.024	-0.969								
ZWC	-0.104	0.113	-0.953								
Median	-0.014	0.024	-0.976								

The table is based on the pre-crisis period  $P_2$ . The OR/RealS and OR/PI correlations are represented by Fisher's  $z$ -transform of correlation coefficients, while RealS/PI correlations are measured using Pearson's correlation coefficient. The critical value for this coefficient is equal to 0.094 at the 5% significance level (436 daily observations). The significant correlation coefficients are marked in italics. Source: authors' calculations.

Table 8. Coefficients of correlation between the values of the daily percentage order ratio, daily percentage realized spread, and daily percentage price impact for 53 WSE-listed companies during the Global Financial Crisis from June 1, 2007 to February 27, 2009

L	OR /RealS	OR/PI	RealS/PI	M	OR /RealS	OR/PI	RealS/PI	S	OR /RealS	OR/PI	RealS/PI
BHW	-0.025	0.022	<i>-0.989</i>	ALM	0.007	-0.020	<i>-0.981</i>	APL	<i>-0.149</i>	<i>0.123</i>	-0.962
BPH	0.056	-0.061	<i>-0.978</i>	AMC	-0.081	0.040	<i>-0.980</i>	BDL	-0.031	-0.009	-0.955
BNP	<i>-0.155</i>	<i>0.158</i>	<i>-0.979</i>	ATG	-0.033	0.034	<i>-0.994</i>	EFK	-0.056	0.046	-0.981
BOS	0.019	-0.019	<i>-0.999</i>	ATM	0.009	-0.013	<i>-0.986</i>	ENP	<i>-0.189</i>	<i>0.153</i>	-0.955
BDX	-0.040	0.026	<i>-0.979</i>	CNG	-0.081	0.067	<i>-0.987</i>	KMP	-0.007	-0.019	-0.984
BZW	<i>0.193</i>	<i>-0.209</i>	<i>-0.985</i>	COL	-0.089	0.068	<i>-0.986</i>	MZA	-0.077	0.071	-0.987
DBC	-0.045	0.032	<i>-0.992</i>	IND	-0.091	0.091	<i>-0.994</i>	PLA	<i>-0.146</i>	<i>0.103</i>	-0.961
ECH	-0.013	-0.004	<i>-0.983</i>	IPL	-0.063	0.058	<i>-0.982</i>	SME	<i>-0.110</i>	<i>0.110</i>	-0.989
GTN	0.007	-0.022	<i>-0.970</i>	LTX	-0.060	0.021	<i>-0.956</i>	Median	-0.094	0.087	-0.971
GTC	0.029	-0.036	<i>-0.971</i>	MCI	-0.100	0.037	<i>-0.963</i>				
ING	-0.094	0.083	<i>-0.982</i>	MNI	-0.053	0.050	<i>-0.963</i>				
KTY	0.023	-0.035	<i>-0.994</i>	PEK	0.015	-0.012	<i>-0.979</i>				
KGH	<i>0.114</i>	<i>-0.143</i>	<i>-0.928</i>	PUE	-0.048	0.073	<i>-0.977</i>				
LPP	-0.023	0.032	<i>-0.982</i>	SKA	0.057	-0.047	<i>-0.995</i>				
MBK	0.002	0.002	<i>-0.978</i>	STF	-0.030	0.019	<i>-0.966</i>				
MIL	0.003	-0.031	<i>-0.987</i>	STX	-0.066	0.034	<i>-0.933</i>				
MOL	0.046	-0.034	<i>-0.990</i>	TIM	-0.056	0.043	<i>-0.988</i>				
NET	-0.047	0.027	<i>-0.983</i>	VST	0.005	-0.029	<i>-0.990</i>				
OPL	0.066	-0.079	<i>-0.977</i>	Median	-0.055	0.036	-0.982				
ORB	0.021	-0.033	<i>-0.995</i>								
PEO	-0.013	0.004	<i>-0.970</i>								
PKN	-0.038	0.040	<i>-0.934</i>								
PKO	-0.068	0.027	<i>-0.934</i>								
STP	-0.040	0.017	<i>-0.987</i>								
SNS	-0.037	0.034	<i>-0.966</i>								
TVN	0.028	-0.053	<i>-0.986</i>								
ZWC	-0.070	0.064	<i>-0.944</i>								
Median	-0.013	0.002	-0.982								

This table is based on the crisis period  $P_3$ . The OR/RealS and OR/PI correlations are represented by Fisher's  $z$ -transform of correlation coefficients, while the RealS/PI correlations are measured using Pearson's correlation coefficient. The critical value of the correlation coefficient is equal to 0.094 at the 5% significance level (436 daily observations). The significant correlation coefficients are marked in italics. Source: authors' calculations.

The results reported in Tables 6–8 are generally consistent with the literature. The majority of the OR/RealS and OR/PI correlation coefficients are not significantly different from zero. On the other hand, all the RealS/PI correlation coefficients are negative and very large as expected, because the proxies for both the realized spread and price impact are treated as effectively components of spread which complement each other.

The strong negative correlation between these two proxies for liquidity confirms that the % RealS (2) and % PI (3) variables are very strongly and negatively associated with each other and capture different sources of market liquidity.

Table 9 summarizes the results of these correlation analyses by presenting the percentage of statistically significant correlation coefficients in the three size groups for all the investigated periods. The evidence reveals that in the case of the large and medium groups the percentage of statistically significant OR/RealS and OR/PI correlations was visibly lower during the crisis period (P<sub>3</sub>) in comparison with other periods.

Table 9. Percentage of statistically significant correlation coefficients

Group	OR/RealS				OR/PI				RealS/PI			
	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>
Large (27 companies)	44.4	33.3	11.1	22.2	48.1	37	11.1	22.2	100			
Medium (18 companies)	88.9	44.4	5.6	38.9	77.8	50	0	22.2				
Small (8 companies)	37.5	62.5	50	37.5	25	37.5	50	12.5				

For explanation, see Table 3. Source: authors' calculations.

## 5. Conclusion

The role of liquidity in empirical finance and the microstructure of markets has grown over the last years influencing conclusions regarding asset pricing, corporate finance, and market efficiency. In his seminal work, Kyle [19] argues that market liquidity is a slippery and elusive concept, in part because it encompasses a number of the transactional properties of markets. For example, the inconsistent evidence of commonality in liquidity on various stock markets all over the world could be attributed to differences between the designs of these markets. It is important to distinguish between order-driven and quote-driven market structures, because market structure determines how orders are transformed into trades and how this transformation affects liquidity. In an order-driven market, no designated market-maker has an obligation to provide liquidity to the market. Traders and investors submit a limit order book to buy and sell shares. Unfortunately, although the WSE is classified as an order-driven market with an electronic order book, information regarding the best bid and ask price is not publicly available. Therefore, various algorithms for inferring the initiator of a trade might help to distinguish between buyer- and seller-initiated trades and they enable us to estimate various proxies for liquidity/illiquidity based on high-frequency intraday data.

Three alternative estimates of liquidity were employed, supported by the Lee-Ready algorithm for inferring the initiator of a trade: (1) the percentage order ratio as an indicator of order imbalance, (2) the percentage realized spread as a temporary component

of the effective spread, and (3) the percentage price impact as a permanent component of the effective spread. The empirical results revealed that the values of all of these proxies for liquidity rather do not depend on a firm's size and turn out to be robust to the choice of the period. Moreover, the correlation coefficients indicate that the proxies for liquidity used in this study seem to capture various sources of market liquidity and therefore might be utilized as liquidity/illiquidity measures in further investigations. Hence, one possibility for continuing this research would be a study on commonality in liquidity on the WSE, because empirical research on the microstructure of markets has recently shifted its focus from the examination of the liquidity of individual securities towards analyses of the common determinants and components of liquidity. Beginning with Chordia et al. [6], the identification of common determinants of liquidity, or commonality in liquidity, has emerged as a new and fast growing strand of the literature on liquidity.

### Acknowledgements

This study was supported by the grant *Comparative research on commonality in liquidity on the Central and Eastern European stock markets* from the National Science Centre in Poland (No. 2016/21/B/HS4/02004).

### References

- [1] AMIHUD Y., MENDELSON H., PEDERSEN L.H., *Liquidity and asset pricing*, Found. Trends Fin., 2005, 1 (4), 269–364.
- [2] BEKAERT G., HARVEY C.R., LUNDBLAD C., *Liquidity and expected returns. Lessons from emerging markets*, Rev. Fin. Stud., 2007, 20 (6), 1783–1831.
- [3] CHAKRABARTY B., LI B., NGUYEN V., VAN NESS R.A., *Trade classification algorithms for electronic communications network trades*, J. Bank. Fin., 2007, 31, 3806–3821.
- [4] CHAKRABARTY B., MOULTON P.C., SHKILKO A., *Short sale, long sale, and the Lee–Ready trade classification algorithm revisited*, J. Fin. Markets, 2012, 15 (4), 467–491.
- [5] CHAN K., FONG W.-M., *Trade size, order imbalance, and the volatility–volume relation*, J. Fin. Econ., 2000, 57, 247–273.
- [6] CHORDIA T., ROLL R., SUBRAHMANYAM A., *Commonality in liquidity*, J. Fin. Econ., 2000, 56, 3–28.
- [7] CHORDIA T., ROLL R., SUBRAHMANYAM A., *Order imbalance, liquidity, and market returns*, J. Finan. Econ., 2002, 65, 111–130.
- [8] CHORDIA T., SARKAR A., SUBRAHMANYAM A., *An empirical analysis of stock and bond market liquidity*, Rev. Finan. Stud., 2005, 18 (1), 85–129.
- [9] COPPEJANS M., DOMOWITZ I., MADHAVAN A., *Resiliency in an automated auction*, Working Paper, 2004.
- [10] FAMA E.F., FRENCH K.R., *Common risk factors in the returns on stocks and bonds*, J. Fin. Econ., 1993, 33 (1), 3–56.
- [11] FISHER R.A., *On the probable error of a coefficient of correlation deduced from a small sample*, Metron., 1921, 1, 3–32.
- [12] GLOSTEN L.R., *Components of the bid-ask spread and the statistical properties of transaction prices*, J. Finance, 1987, 42 (4), 1293–1307.
- [13] GOYENKO R.Y., HOLDEN C.W., TRZCINKA C.A., *Do liquidity measures measure liquidity?*, J. Fin. Econ., 2009, 92, 153–181.
- [14] HAMEED A., KANG W., VISWANATHAN S., *Stock market decline and liquidity*, J. Fin., 2010, 65 (1), 257–293.



- [15] HUANG R.D., STOLL H.R., *Dealer versus auction markets. A paired comparison of execution costs on NASDAQ and the NYSE*, J. Fin. Econ., 1996, 41, 313–357.
- [16] HUANG R.D., STOLL H.R., *The components of the bid-ask spread. A general approach*, Rev. Fin. Stud., 1997, 10 (4), 995–1034.
- [17] HUBERMAN G., HALKA D., *Systematic liquidity*, J. Fin. Research, 2001, 24 (2), 161–178.
- [18] KORAJCZYK R., SADKA R., *Pricing the commonality across alternative measures of liquidity*, J. Fin. Econ., 2008, 87 (1), 45–72.
- [19] KYLE A.S., *Continuous auctions and insider trading*, Econometrica, 1985, 53 (6), 1315–1336.
- [20] LEE C.M.C., READY M.J., *Inferring trade direction from intraday data*, J. Fin., 1991, 46 (2), 733–746.
- [21] LESMOND D.A., *Liquidity of emerging markets*, J. Fin. Econ., 2005, 77 (2), 411–452.
- [22] LESMOND D.A., OGDEN J.P., TRZCINKA C.A., *A new estimate of transaction costs*, Rev. Fin. Stud., 1999, 12 (5), 1113–1141.
- [23] NOWAK S., *Order imbalance indicators in asset pricing. Evidence from the Warsaw Stock Exchange*, [in:] K. Jajuga, L. Orłowski, K. Staehr (Eds.), *Contemporary Trends and Challenges in Finances*, Springer Proc. in Business and Economics, 2017, 91–102.
- [24] NOWAK S., OLBRYŚ J., *Direct evidence of non-trading on the Warsaw Stock Exchange*, Research Papers of Wrocław University of Economics, 2016, 428, 184–194.
- [25] OLBRYŚ J., *Capital Asset Pricing on a Market with Frictions in Trading Processes*, Difin Press, Warsaw 2014 (in Polish).
- [26] OLBRYŚ J., *Is illiquidity risk priced? The case of the Polish medium-size emerging stock market*, Bank i Kredyt, 2014, 45 (6), 513–536.
- [27] OLBRYŚ J., *Interaction between market depth and market tightness on the Warsaw Stock Exchange. A preliminary study*, [in:] K. Jajuga, L. Orłowski, K. Staehr (Eds.), *Contemporary Trends and Challenges in Finances*, Springer Proc. Business and Economics, 2017, 103–111.
- [28] OLBRYŚ J., MAJEWSKA E., *Direct identification of crisis periods on the CEE stock markets. The influence of the 2007 U.S. subprime crisis*, Proc. Econ. Fin., 2014, 14, 461–470.
- [29] OLBRYŚ J., MAJEWSKA E., *Bear market periods during the 2007–2009 financial crisis. Direct evidence from the Visegrad countries*, Acta Oecon., 2015, 65 (4), 547–565.
- [30] OLBRYŚ J., MURSZTYN M., *Comparison of selected trade classification algorithms on the Warsaw Stock Exchange*, Adv. Comp. Sci. Res., 2015, 12, 37–52.
- [31] OLBRYŚ J., MURSZTYN M., *Dimensions of market liquidity. The case of the Polish stock market*, [in:] N. Tsounis, A. Vlachvei (Eds.), *Advances in Applied Economic Research*, Springer Proc. Business Econ., 2017, 151–166.
- [32] PAGAN A.R., SOSSOUNOV K.A., *A simple framework for analysing bull and bear markets*, J. Appl. Econ., 2003, 18 (1), 23–46.
- [33] RANALDO A., *Intraday market liquidity on the Swiss Stock Exchange*, Swiss Society for Financial Market Research, 2001, 15 (3), 309–327.
- [34] ROLL R., *A simple implicit measure of the effective bid-ask spread in an efficient market*, J. Fin., 1984, 39 (4), 1127–1140.
- [35] STOLL H.S., *Friction*, J. Fin., 2000, 55 (4), 1479–1514.
- [36] THEISSEN E., *A test of the accuracy of the Lee/Ready trade classification algorithm*, J. Int. Fin. Mark., Inst. Money, 2001, 11, 147–165.
- [37] TSAY R.S., *Analysis of Financial Time Series*, Wiley, New York 2010.
- [38] WYSS R. VON, *Measuring and predicting liquidity in the stock market*, Dissertation No. 2899, University of St. Gallen, 2004.

Received 19 May 2017

Accepted 12 December 2017