HIGHER SCHOOL OF ECONOMICS
NATIONAL RESEARCH UNIVERSITY

Dmitry Borisenko, Sergey Gelman

LIQUIDITY, ASYMMETRIC INFORMATION AND ASSET PRICING ON THE RUSSIAN STOCK MARKET

Working Paper WP9/2012/01
Series WP9
Research of economics
and finance

Moscow
2012

Editor of the Series WP9
“Research of economics and finance”
A. Belianin

Borisenko, D., Gelman, S.* Liquidity, asymmetric information and asset pricing on the Russian stock market** : Working paper WP9/2012/01 [Text] / D. Borisenko, S. Gelman ; The University – Higher School of Economics. – Moscow : Publishing House of the University – Higher School of Economics, 2012.

In this paper we investigate how asymmetric information and informed trading influences liquidity and how liquidity influences asset pricing on the Russian stock market in 1998–2011. We use a battery of existing liquidity proxies as well as our own modification of Lesmond et al. (1999) measure and capture informed trading through positive daily return autocorrelation. We find that asymmetric information worsens liquidity, whereas no supportive evidence of adverse impact of financial distress and informed trading can be discovered, which could be in the latter case partly due to a weak proxy. Furthermore, liquidity, along with market risk, seems to be the major driver of asset pricing on the Russian stock market. This result, however, is not robust to specifying liquidity as characteristic rather than factor.

JEL-Classification: G12, G14.

Dmitry Borisenko
International College of Economics and Finance
National Research University Higher School of Economics
Pokrovskij Bulvar 11, 109028 Moscow, Russia

Sergey Gelman
International College of Economics and Finance
National Research University Higher School of Economics
Pokrovskij Bulvar 11, 109028 Moscow, Russia
Email: sgelman@hse.ru

* Corresponding author.

** We would like to thank the participants of ICEF mini-conference on financial markets for helpful comments and Artem Borisenko for excellent research assistance. Sergey Gelman gratefully acknowledges the financial support of the NRU HSE Scientific Fund, project # 10-01-0110.

1. Introduction

Liquidity is a main concern of a stock exchange, as well as of participating market agents. Russian and foreign investors could be possibly underinvesting in Russian stocks if they deem liquidity as insufficient. In light of establishing Moscow as an international financial center (which has been declared by the federal government as political objective for the nearest future), finding means of increasing liquidity becomes even more relevant. The major Russian stock exchange, Moscow Interbank Currency Exchange, having a rather technologically top notch trading system, which allows very cheap and quick small order execution, acknowledges itself an insufficient depth of the market, which led it to suggest a “dark pool” in 2011.

In this paper, we empirically investigate theoretical predictions for the interplay of liquidity, liquidity provision and asset pricing using daily stock prices from the Moscow Interbank Currency Exchange from the period 1998–2011. The data are especially insightful for studying the link between insider behavior and liquidity because Russian insider legislation was introduced only in 2011 and massive insider trading on the market can be suspected. Moreover, anecdotal evidence of intensive stock market involvement of oligarchs, high-ranked government officials and other informed players regularly leaks into mass-media.

We use theoretical predictions based on stock return dynamics to track down informed trading: a high negative serial correlation of stock returns indicates high bid-ask spreads (Roll 1984), whereas a positive correlation of returns could indicate speculation based on private information of a large stakeholder (Llorente et al. 2002).

For the analysis of liquidity we exploit several existing liquidity measures as well as suggest a new measure of transaction costs controlling for relatively high idiosyncratic volatility.

Our results provide supportive evidence of the asymmetry theory (Kyle 1985, Rochet and Vila 1994): size is a strong driver of liquidity, significant in nearly all specifications for all measures. Since we control for the trading volume channel of size influence on liquidity by including trading activity, thus the effect of size can be mostly explained by better information coverage of large companies. Additional one hundred billion rubles of market capitalization reduces round-trip transaction costs by 2–2.5 percentage points. Trading activity enhances liquidity significantly either.

However our data does not provide supportive evidence of the negative impact of informed trading on liquidity: autocorrelation coefficient is in majority of cases insignificant, and in some exceptional cases significant, but with a contrary sign. Possibly, either investors react instantly on insider trading and liquidity goes down in the same period, thus there are no remaining dynamical effects, or autocorrelation coefficient is an inadequate proxy of insider trading in case of MICEX.

We use both portfolios and individual stocks to test the pricing effect of liquidity, both as characteristic and as factor. We find supportive evidence of the positive influence of liquidity on asset pricing, however rather as factor, not as characteristic: economical significance of the impact of transaction costs on expected returns is negligible, whereas the liquidity risk premium is about 18% on the annual basis (and statistically significant on the 1% level). Market risk and liquidity risk seem to be the only relevant factors on the Russian stock market: when extending the model with these two factors to encompass the Carhart's (1997) model, the remaining size, book-to-market and momentum factors turn out to be insignificant.

The rest of the paper is structured as follows: Section 2 describes the trading microstructure of Moscow Interbank Currency Exchange and give some background information, section 3 provides an overview of liquidity measures and develops a new measure of liquidity, section 4 describes the data and section 5 treats main econometric techniques, used in the empirical study. Results of analysis of determinants of liquidity and of impact of liquidity on asset pricing are presented and discussed in section 6. Section 7 concludes.

2. Market structure and historical background

Stock trading in Russia in the period under study was concentrated on three organized markets: Russian Trading System (RTS), Sankt-Petersburg Stock Exchange (SPbSE) and Moscow Interbank Currency Exchange (MICEX). MICEX established its stock section in May 1997 and rapidly exceeded other Russian exchanges in terms of total stock turnover. Figures 1 and 2 summarize the dynamics of total turnover of the Russian stock market. In 1999 more than 57% of all organized stock transactions (by value) were executed on MICEX. During the subsequent years MICEX efficiently represented the whole market.

Figure 1. Total year turnover on MICEX and other exchanges 1995–1999

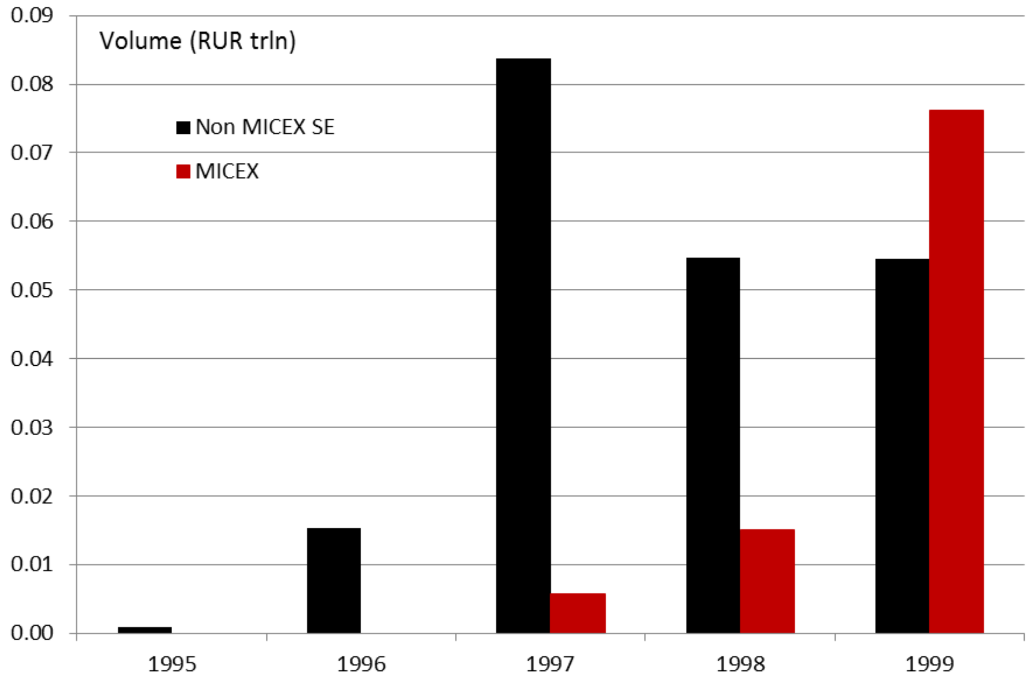
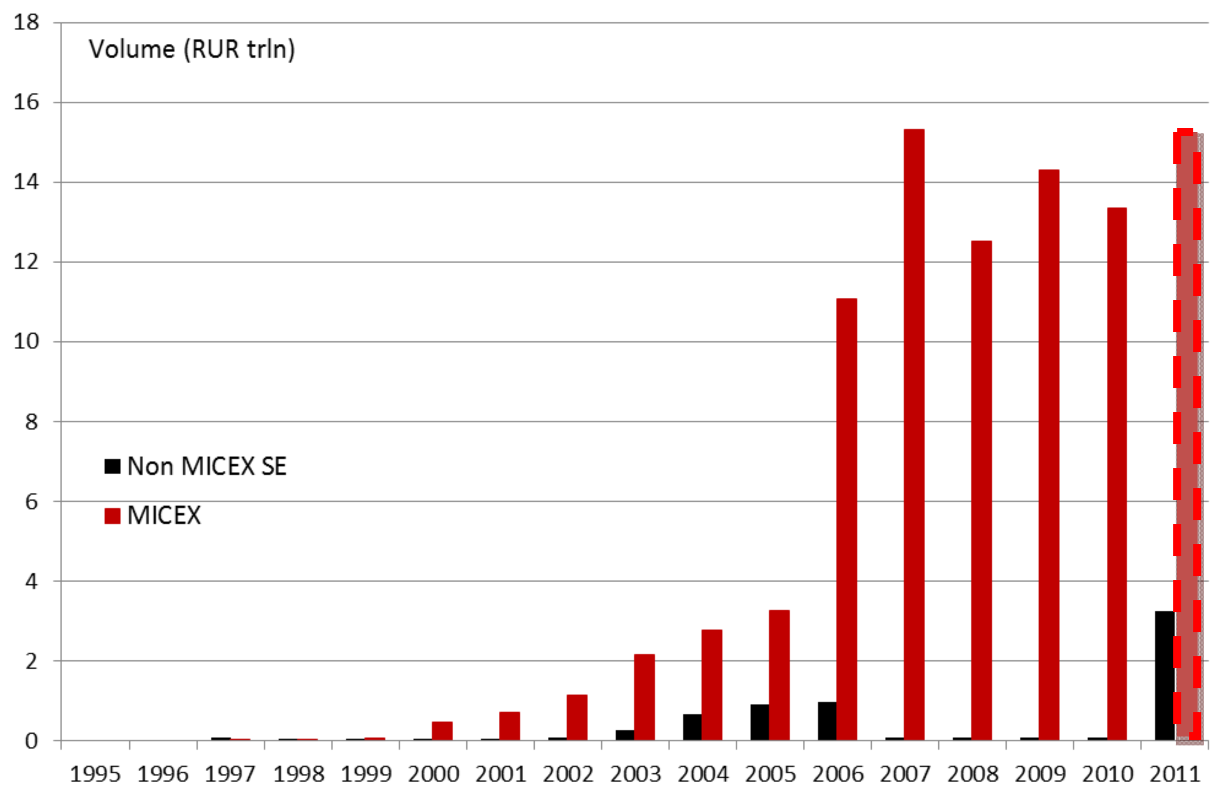


Figure 2. Total year turnover on MICEX and other exchanges 1995–2011



Source: RTS-MICEX statistics, own estimates (for MICEX 2011).

MICEX is an electronic stock exchange with continuous trading. Major trading form is continuous auction, whereas electronic order book to a large extent (best 10 orders on both sides) is open to all participating brokers. Both limit orders and market maker orders are allowed. All orders are submitted electronically.

On top of matching all incoming orders MICEX provides additional liquidity through market makers and specialists. MICEX participants may obtain status of market-maker for securities with turnover less than 100 mln rubles per month. According to the Trading Rules a market-maker can simultaneously set buy and sell quotes, but the spread between the quotes should be below a threshold, set by the exchange:

$$Actual\ Spread = \frac{P_1 - P_2}{P_2} \leq Marginal\ Spread$$

where P_1 is the maximum price of sell orders and P_2 is the minimum price of buy orders submitted by the market maker, and *marginal spread* is determined by the Exchange for each stock. Market-maker has the right to submit orders only in one direction, in such case deviation from the previously quoted price should not exceed the half of marginal spread. Prior to 2005 market-makers were also obliged to provide minimum turnover specified by the Exchange Board for each security. In July 2005 the turnover restriction for market-makers was abolished. In addition in December 2005 specialists were introduced. In contrast to market-makers, specialists are able to submit all types of orders including orders addressed to a particular counterparty.

Stocks are generally traded in standard lots of 100 shares; however it is possible to initiate transactions for a lower number of stocks. The trading day starts with an opening auction, where the price is set at 10:30. Thereafter continuous trading session runs until 18:45. After the closing of the session traders still can submit market orders, which are executed at 19:00 at a weighted average price of the last 30 minutes of the trading session.

MICEX retains a commission from 0.00004 to 0.00007 of the value of transaction, but not less than 0.06 rubles (Clearing rules, 2010).

3. Liquidity measures

We use several existing liquidity measures which are described in this subsection.

Bid-Ask Spread and Commissions

Spread is determined by market-makers who quote the range of prices at which transactions may be immediately executed or by specialists who submit orders in opposite direction to order flow from market participants. Commissions are taken by exchanges and brokers usually as a fraction of transaction value (MICEX commission is 0.004% to 0.007%, see Section 2).

$$Spread = \frac{P_{ask} - P_{bid}}{P} \tag{1}$$

Amihud Measure

Amihud (2002) defines liquidity measure as average ratio of absolute return to the total turnover:

$$Amihud_{jt} = \frac{|R_{jt}|}{P_{jt} Volume_{jt}} \tag{2}$$

the absolute percentage price impact of a unit of trading volume. This measure represents the ‘depth’ of the market or price sensitivity to order flow. The Amihud estimates are usually multiplied by 10^6 in order to get a scale comparable to other measures.

LOT Measure

An informed investor will trade only if the value of information he or she possesses exceeds the transaction costs. Lesmond et al. (1999) propose the measure that extracts transaction costs from the return data. The LOT measure includes not only spread plus commission but also it reflects the other costs met by trader. For example, costs associated with information acquisition, opportunity costs, expected price impact, etc. This approach suggests that there is no change in price whenever the unobservable ‘true’ price does not exceed costs threshold. The model is set up as follows:

$$\begin{aligned}
R^*_{jt} &= \beta_j R_{mt} + \mathrm{T}_{jt} \\
R_{jt} &= R^*_{jt} - \alpha_{1j}, \ \ if \qquad R^*_{jt} < \alpha_{1j} \\
R_{jt} &= 0, \qquad\quad if \ \alpha_{1j} \le R^*_{jt} \le \alpha_{2j} \\
R_{jt} &= R^*_{jt} - \alpha_{2j}, \ if \qquad R^*_{jt} > \alpha_{2j}
\end{aligned} \tag{3}$$

Where R_{jt}^{*} (the true return of firm j) follows the market model with suppressed intercept, and α_{1j}, α_{2j} are the costs of selling and buying respectively. The difference $\alpha_{2j} - \alpha_{1j}$ is the measure of round-trip transaction costs. The model is estimated by maximizing the following likelihood function:

$$L(\alpha_{1j}, \alpha_{2j}, \beta_j, \sigma_j \mid R_{jt}, R_{mt}) = \\ = \prod_{\Omega_0} \left(\Phi_{2j} - \Phi_{1j} \right) \times \prod_{\Omega_1} \frac{1}{\sigma_j} \phi \left(R_{jt} + \alpha_{1j} - \beta_j R_{mt} \right)^2 \prod_{\Omega_2} \frac{1}{\sigma_j} \phi \left(R_{jt} + \alpha_{2j} - \beta_j R_{mt} \right)^2 \quad (4)$$

The corresponding log-likelihood is:

$$LogL(\alpha_{1j}, \alpha_{2j}, \beta_j, \sigma_j \mid R_{jt}, R_{mt}) = \sum_{\Omega_0} ln \left[\Phi_{2j} \left(\frac{\alpha_{2j} - \beta_j R_{mt}}{\sigma_j} \right) - \Phi_{1j} \left(\frac{\alpha_{1j} - \beta_j R_{mt}}{\sigma_j} \right) \right] + \\ + \sum_{\Omega_1} ln \frac{1}{\left(2\pi\sigma_j^2 \right)^{\frac{1}{2}}} - \sum_{\Omega_1} \frac{1}{2\sigma_j^2} \left(R_{jt} + \alpha_{1j} - \beta_j R_{mt} \right)^2 + \sum_{\Omega_2} ln \frac{1}{\left(2\pi\sigma_j^2 \right)^{\frac{1}{2}}} - \sum_{\Omega_2} \frac{1}{2\sigma_j^2} \left(R_{jt} + \alpha_{2j} - \beta_j R_{mt} \right)^2 \quad (5)$$

Where Ω_1, Ω_2 are the regions of nonzero returns with negative and positive market return respectively, Φ, ϕ are the standard normal distribution and density functions.

3.1. New measure and its properties

The LOT model could be biased upwards in case of relatively high idiosyncratic return variance. It treats observed counter-market stock price movements, which are due to idiosyncratic shocks, as very high transaction costs. I. e. LOT regions Ω_1, Ω_2 are defined over positive and negative market returns respectively, however one may observe negative stock returns on growing market. These returns are caused by significant negative shocks which a.) override the information provided by the market and b.) exceed the sell costs threshold α_{1j}. The same intuition is applied for positive stock returns on falling market. While LOT attempts to estimate costs irrespectively of the direction of transactions and allows for the abovementioned mistreatment, we develop the model that treats returns made on market information and idiosyncratic shocks separately. Stock j return is defined over 5 regions:

$$
\begin{aligned}
&R_{jt}^{*} = \beta_j R_{mt} + \tau_{jt} \\
&R_{jt} = R_{jt}^{*} - \alpha_{1j}, \quad for\ \Lambda_1 = \{If\ R_{mt} < 0\ and\ R_{jt} < 0\} \\
&R_{jt} = R_{jt}^{*} - \alpha_{2j}, \quad for\ \Lambda_2 = \{If\ R_{mt} > 0\ and\ R_{jt} > 0\} \\
&R_{jt} = R_{jt}^{*} - \alpha_{1j}, \quad for\ \Lambda_3 = \{If\ R_{mt} > 0\ and\ R_{jt} < 0\} \\
&R_{jt} = R_{jt}^{*} - \alpha_{2j}, \quad for\ \Lambda_4 = \{If\ R_{mt} < 0\ and\ R_{jt} > 0\} \\
&R_{jt} = 0, \qquad\qquad\ \ for\ \Lambda_0 = \{If\ \alpha_{1j} < R_{jt}^{*} < \alpha_{2j}\}
\end{aligned}
\tag{6}
$$

By construction the residual term across regions Λ_3 and Λ_4 is not normally distributed. It rather represents the right and left tails of the distribution. Maintaining the assumption of idiosyncratic shocks we employ the Fisher-Tippet-Gnedenko theorem to make an inference about the distribution:

Theorem: Let $(X_1, X_2, ..., X_n)$ be a sequence of i.i.d. random variables, and $M_n = \max\{X_1, X_2, ..., X_n\}$. If $\exists(a_n, b_n): a_n > 0$, and $\lim_{n\to\infty} P\left((M_n - b_n)/a_n \le x\right) = F(x)$, and $F(x)$ is a non-degenerate distribution, then it belongs to the Generalized Extreme Value distribution (either the Gumbel, the Frechet, or the Weibull family.)

We utilize the type I extreme value (or the Gumbel) distribution for maxima in region Λ_4 and for minima in region Λ_3. The probability density functions for these distributions are:

$$
f_{max}(\varepsilon_4) = \frac{1}{s} exp\left\{\frac{-\varepsilon_4}{s} - exp\left(\frac{-\varepsilon_4}{s}\right)\right\}, \qquad \text{and} \qquad f_{min}(\varepsilon_3) = \frac{1}{s} exp\left\{\frac{\varepsilon_3}{s} - exp\left(\frac{\varepsilon_3}{s}\right)\right\}
$$

where $\varepsilon_{3,4}$ are the residuals from the corresponding equations of system (6). Note that the shape parameter s is functionally linked to the standard deviation of residuals: $s = \sigma_\varepsilon \sqrt{6} / \pi$.

The resulting likelihood function for the Generalized Extreme Value (GEV) limited dependent variable model is:

$$
\begin{aligned}
L(\alpha_{1j}, \alpha_{2j}, \beta_j, \sigma_j \mid R_{jt}, R_{mt}) = &\prod_{\Lambda_0} \left(\Phi_{2j} - \Phi_{1j}\right) \times \\
&\times \prod_{\Lambda_1} \frac{1}{\sigma_j} \phi\left(R_{jt} + \alpha_{1j} - \beta_j R_{mt}\right)^2 \prod_{\Lambda_2} \frac{1}{\sigma_j} \phi\left(R_{jt} + \alpha_{2j} - \beta_j R_{mt}\right)^2 \times \\
&\times \prod_{\Lambda_3} f_{min}\left(R_{jt} + \alpha_{1j} - \beta_j R_{mt}\right) \prod_{\Lambda_4} f_{max}\left(R_{jt} + \alpha_{2j} - \beta_j R_{mt}\right)
\end{aligned}
\tag{7}
$$

The Log-likelihood is:

$$LogL(\alpha_{1j},\alpha_{2j},\beta_j,\sigma_j \mid R_{jt},R_{mt}) = \sum_{\Lambda_0} ln\left[\Phi_{2j}\left(\frac{\alpha_{2j}-\beta_j R_{mt}}{\sigma_j}\right)-\Phi_{1j}\left(\frac{\alpha_{1j}-\beta_j R_{mt}}{\sigma_j}\right)\right]+$$
$$+\sum_{\Lambda_1} ln\frac{1}{\left(2\pi\sigma_j^2\right)^{\frac{1}{2}}}-\sum_{\Lambda_1}\frac{1}{2\sigma_j^2}\left(R_{jt}+\alpha_{1j}-\beta_j R_{mt}\right)^2+\sum_{\Lambda_2} ln\frac{1}{\left(2\pi\sigma_j^2\right)^{\frac{1}{2}}}-\sum_{\Lambda_2}\frac{1}{2\sigma_j^2}\left(R_{jt}+\alpha_{2j}-\beta_j R_{mt}\right)^2+ \quad (8)$$
$$+\sum_{\Lambda_3} ln\left(f_{min}\left(R_{jt}+\alpha_{1j}-\beta_j R_{mt}\right)\right)+\sum_{\Lambda_4} ln\left(f_{max}\left(R_{jt}+\alpha_{2j}-\beta_j R_{mt}\right)\right)$$

The Gumbel distribution CDF is logarithmically concave, hence optimization of the equation (8) yields the global maximum.

3.2. LOT and GEV Comparison: A Simulation Study

In order to compare LOT and GEV measures we perform the simulations study. The experiments are generally structured as follows:

1) Market return series follow a constant mean process with normally distributed disturbance term.

2) True return R_{jt}^* follows the market model with normally distributed disturbance term and suppressed intercept.

3) R_{jt}^* that lie within the exogenously given transaction costs interval are set to zero.

4.) Both models are estimated on 1000 generated series

5) One of the parameters such as return noise variance is gradually changed and new series are generated.

The standard errors of the round-trip costs estimates were evaluated from coefficient variance-covariance matrices obtained during maximization of log-likelihoods (5), (8):

$$S.E. = \sqrt{\text{var}\left(\alpha_{2j}-\alpha_{1j}\right)} = \sqrt{\left(\text{var}\left(\alpha_{2j}\right)-2\,\text{cov}\left(\alpha_{2j},\alpha_{1j}\right)+\text{var}\left(\alpha_{1j}\right)\right)} \quad (9)$$

Experiment 1. Transaction Costs Threshold

The results for growing transaction costs threshold are depicted in Figure 3. LOT estimates exhibit four-fold positive bias when transaction costs are low. As costs increase bias reduces and ultimately becomes negative, thus the true value belongs to 95% confidence interval. The bias is attributed to the fact that LOT explicitly estimates buy costs in the region where sell transactions with non-zero return were observed. Figure 4 illustrates the results for GEV estimates which appear to be very precise.

Experiment 2. Robustness Check. Idiosyncratic stock return Variance

Figures 5 and 6 report the robustness check for LOT and GEV respectively, namely we investigate how the estimates are affected by idiosyncratic stock return variance. As one can notice the GEV estimates remain unbiased when the return variance is increased three-fold from the original value, while LOT experiences monotonic increase in bias.

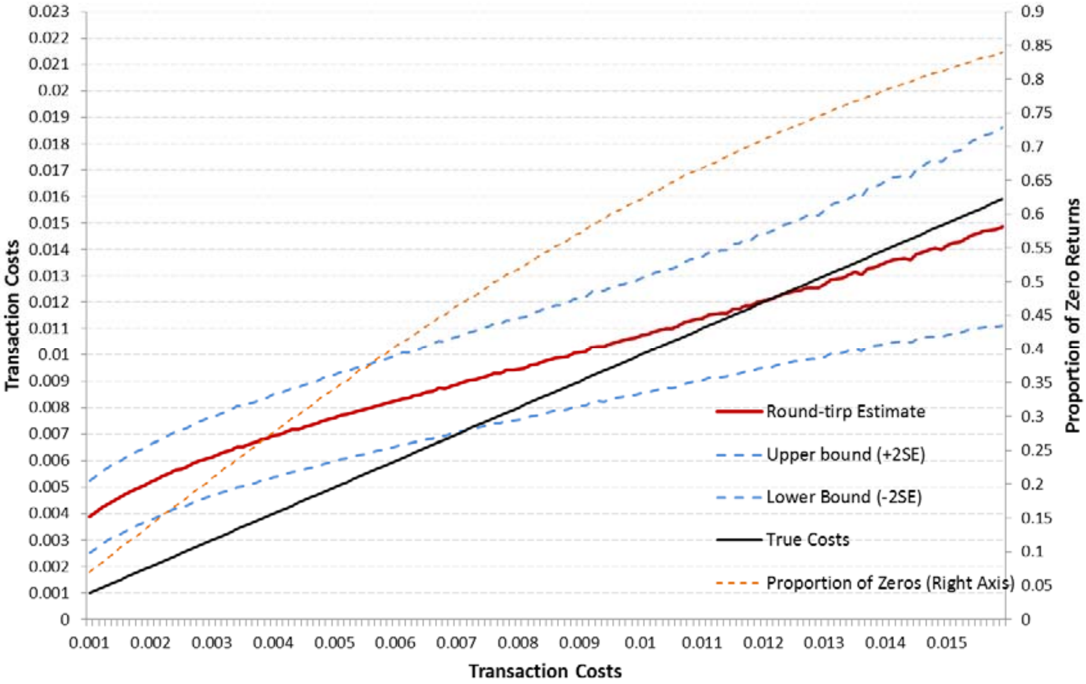
Experiments 3 and 4. Robustness Check. Beta and Market Variance

The robustness check for beta does not yield any change in bias for LOT or bias for GEV. Figures A1 and A2 provide evidence that the Lesmond et al. (1999) model still yields a significant and persistent bias, while GEV estimates are always within the confidence interval and are very close to the true transaction costs value. All of the above also applies to the market return variance (results are reported in Figures A3 and A4).

Experiment 5. Robustness Check. Altering the Disturbance Distribution

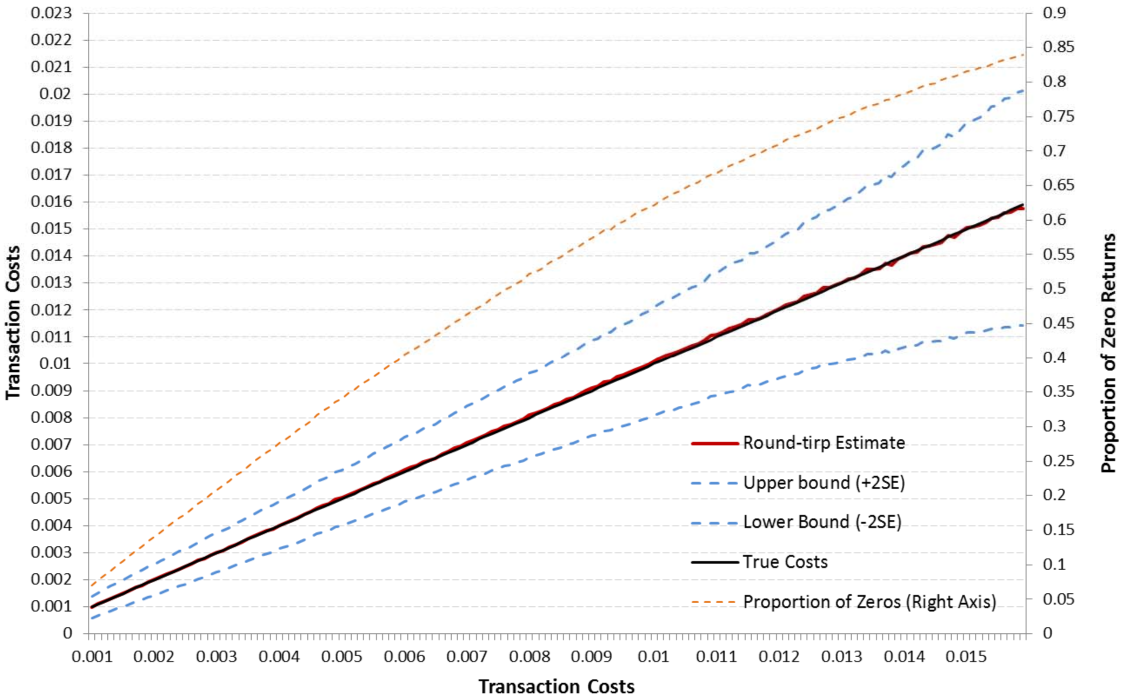
The previous experiments indicated that only parameters that seem to affect the estimates of both LOT and GEV models are the return variance and the transaction costs threshold. The current series of experiments examines the robustness of estimates when the error distribution departs from the normal distribution. In order to capture leptokurtic distributions often observed on the real data we utilize Student's t-distribution with 5, 8, 10, and 12 degrees of freedom. As one can see from the Figures 7, 8, 9, and 10 the GEV model performs better until the true costs exceed 0.013 which corresponds to approximately 73% of zero return days for distribution with 12 degrees of freedom; 0.0135 (74% of zeros) for 10 degrees of freedom; and from 0.014 (75%) and 0.0155 (78%) for distributions with 8 and 5 degrees of freedom respectively.

Figure 3. Simulation results: LOT estimates vs. true transaction costs at different transaction costs levels



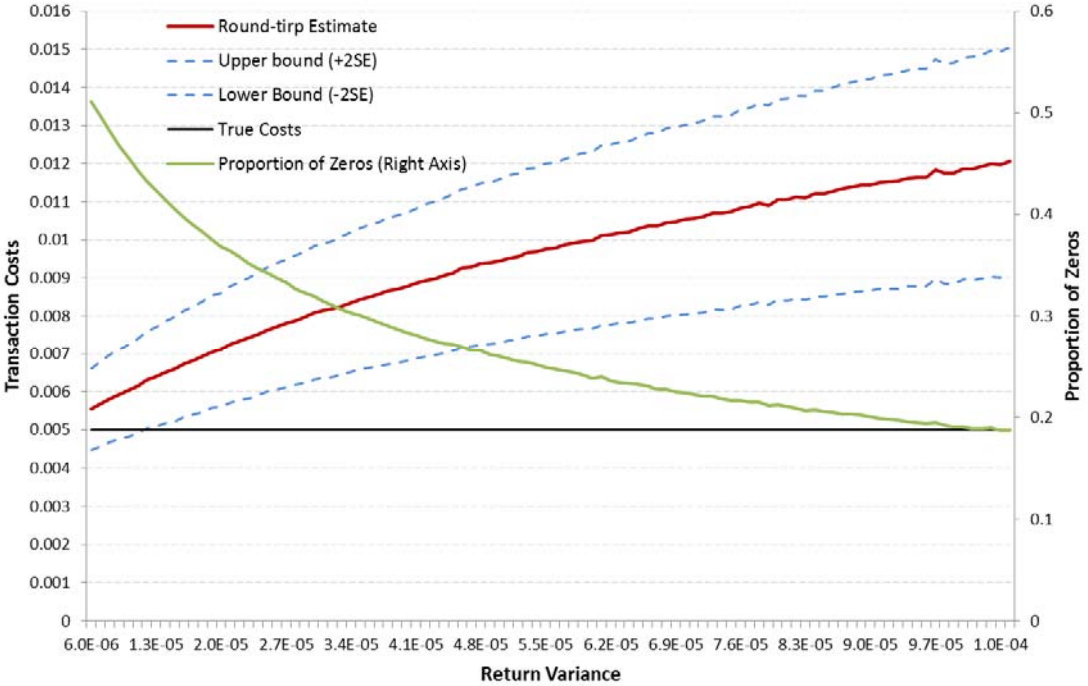
The latent ('true') return follows the process $R_t^* = 0.92R_{mt} + 0.05^2 z_t, z_t \sim N(0,1)$, $R_{mt} \sim N\left(2.5\times10^{-4}, 8.3\times10^{-6}\right)$. Solid black line represents exogenously given bounds (true costs), for each data point 1000 series were generated and lines represent the mean across 1000 estimates. Dashed orange line is the proportion of zero return days (scaled on the right axis). Each estimation cycle from 1 to 150 the cost threshold was increased by 10^{-4} starting from 0.001 and ending at 0.0159, so the black solid line actually has a 45 degree slope. The diagram reads as follows: when the actual round-trip costs equal 0.01 the average estimate is 0.0108, and there are 63% zero return days. The estimation sample includes 350 observations.

Figure 4. Simulation results: GEV estimates vs. true transaction costs at different transaction costs levels



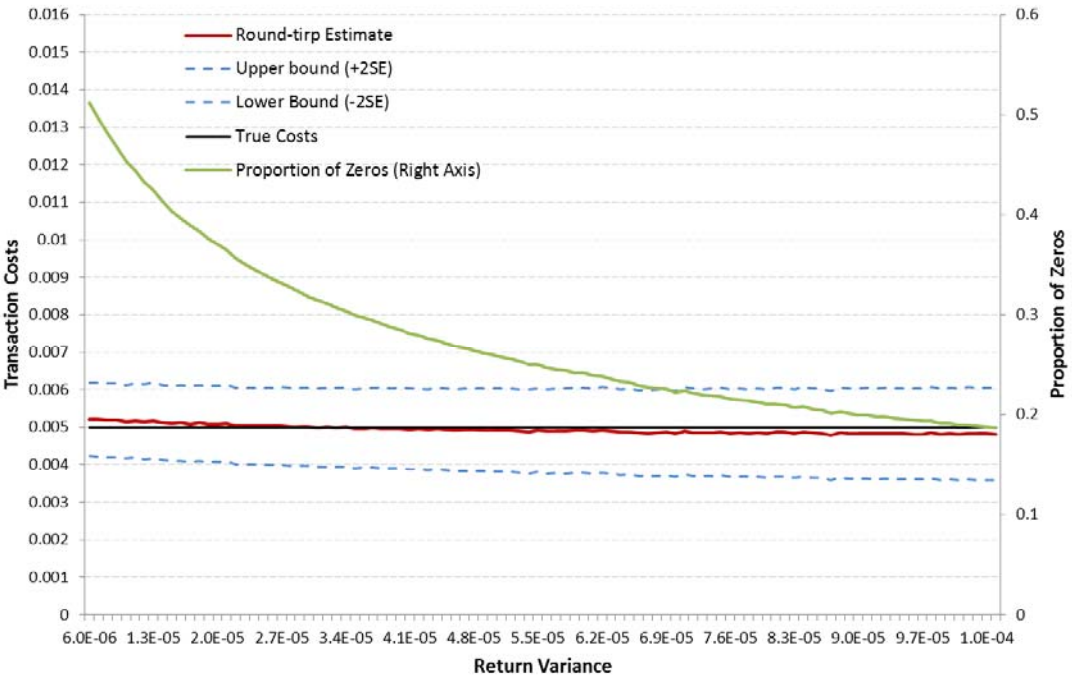
Reads the same as Figure 1.

Figure 5. Simulation results: LOT estimates vs. true transaction costs at different idiosyncratic latent return variance levels



The green solid line represents the proportion of zero returns (scaled on the right axis).Each simulation step return variance increased from 6×10^{-6} to 1.05×10^{-4} with increment of 10^{-6}. So there is a total of 100 steps, each estimating the model 1000 times for 350 observations.

Figure 6. Simulation results: GEV estimates vs. true transaction costs at different idiosyncratic latent return variance levels



Reads similar to Figure 3.

Figure 7. GEV vs. LOT estimates at different transaction costs levels, t-dist. with 12 d.f.

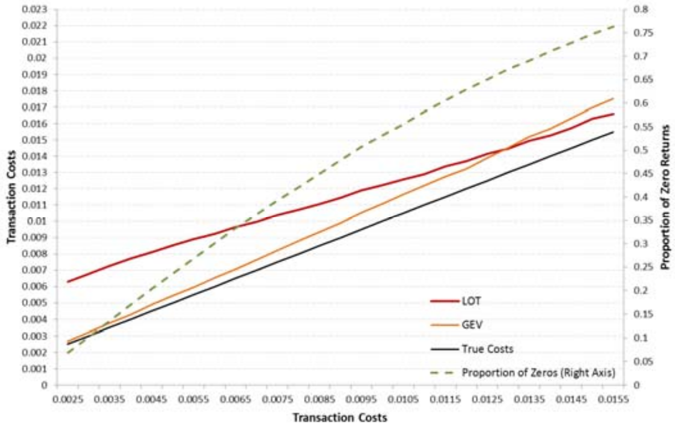


Figure 8. GEV vs. LOT estimates at different transaction costs levels, t-dist. with 10 d.f.

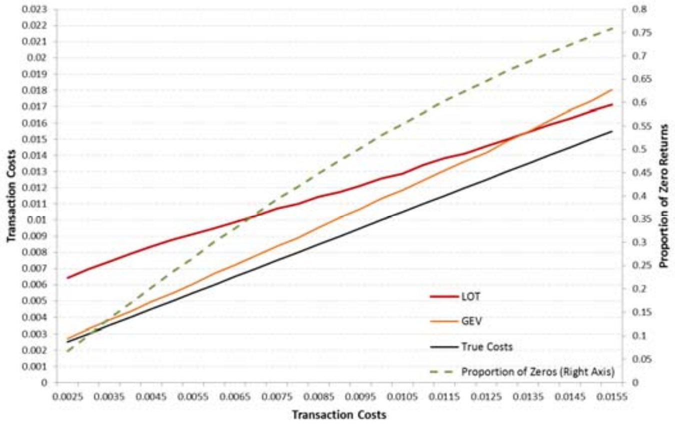


Figure 9. GEV vs. LOT estimates at different transaction costs levels, t-dist. with 8 d.f.

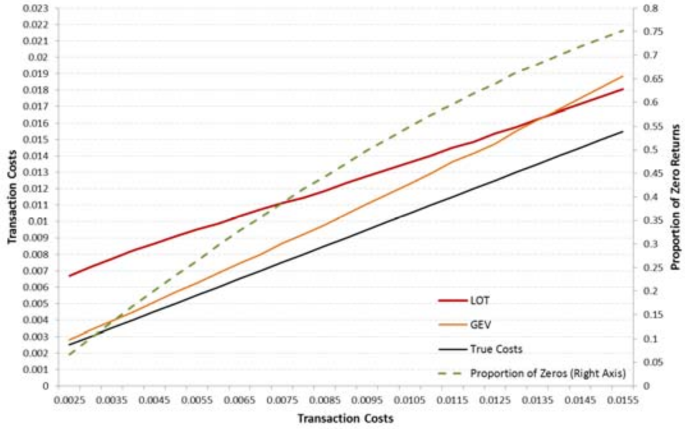
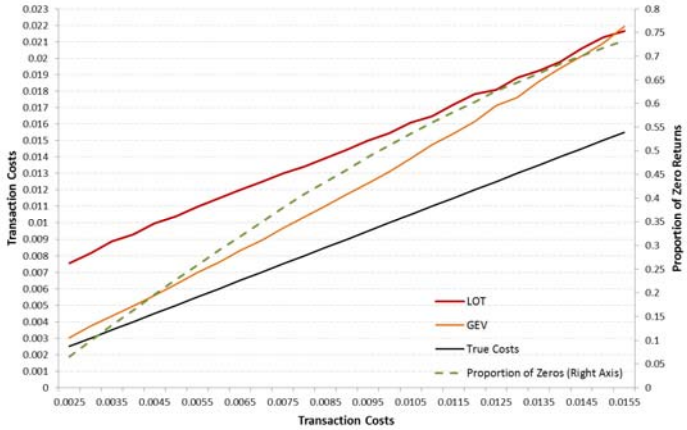


Figure 10. GEV vs. LOT estimates at different transaction costs levels, t-dist. with 5 d.f.



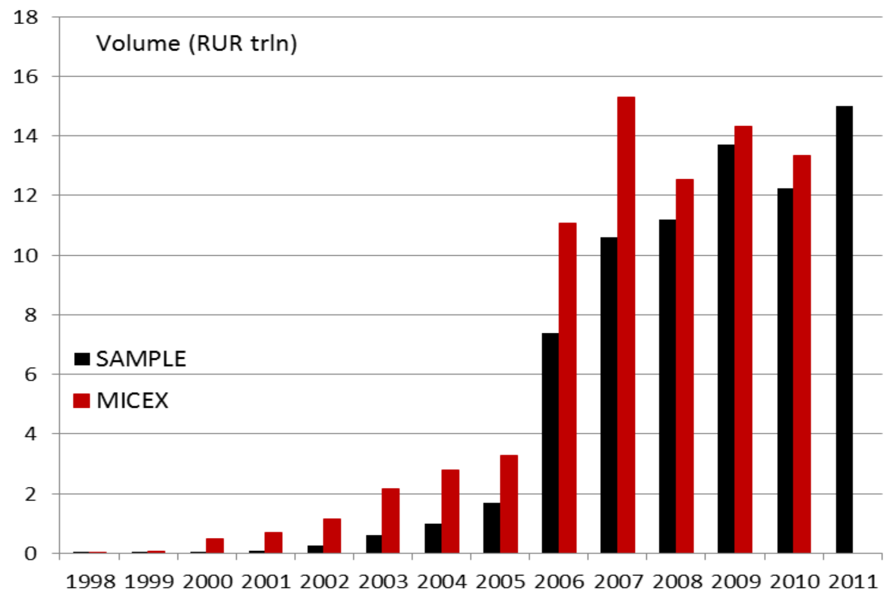
The return process is similar to experiment 1, except that the disturbance term is *t*-distributed with the number of degrees of freedom reflected in the figure title.The step in transaction costs threshold is now increased from 0.0001 to 0.0005. So there is a total of 30 repetitions for each distribution choice.

4. Data and descriptive statistics

We obtain daily closing prices and trading volumes of 310 stocks, traded on MICEX in 2011 from financial information agency Finam. The period covered starts on January 1, 1998 and ends on December 31, 2011. For stocks, first time listed on MICEX after January 1, first day corresponds to the first trading day on MICEX. Closing bid and ask prices for the same period are obtained from Datastream. Also balance sheet type data, such as market value, market to book ratio, free float value and dividend yield are from Datastream.

Judging by turnover value (see Figure 11) our sample covers more then a half of trading activity in MICEX stocks from 2005 onwards (up to 95% in 2010). Prior to 2005 it still delivers a representative draw of Russian stocks.

Figure 11. MICEX and sample year turnovers 1998–2011



We construct the primary direct liquidity proxy “quoted spread” subtracting the bid price from the ask price of the stock at the same date and dividing the difference by the reported closing price. We drop the dates where either bid or ask price is not reported by Datastream. Moreover we also drop observations where the spread turns out to be negative, as it could indicate stale quotes or reporting errors (however, there are only 63 such observations out of more than 240 thousand, so it barely influences the results). Spreads higher than 3 (300%) are set to be equal to 3.

To calculate average spread for each year we take an average of all observed daily spreads for the corresponding period. One should note, that quoted spread reflects transaction costs in continuous trading, which could be higher than those in the opening auction or in the post-session trading and could be interpreted as turnaround cost of immediate execution (of a buy and a sell) before commissions.

We calculate annual averages of Amihud measure, whereby daily Amihud measures are capped at 10. We also calculate annual values of LOT measure and suggested in this paper GEV measure according to the methodology outlined in Section III.

We drop 22 stocks from the sample, which have less than 100 nonzero volume trading days during our 14 year period, reducing the sample to 288 companies. To get reasonable comparison basis we have also dropped stock-year observations with LOT and GEV measures above 3. Find below descriptive statistics for the set of stocks under study.

Table 1. Descriptive statistics for the full period, 1998–2011

	Volume (mill. rubles)	**Market Value (bill. rubles)**	**Free Float Value (bill. rubles)**	**Market-to-Book ratio**
Mean	29100	69.2	57.4	2.36
Median	156	4.4	4.4	1.36
St. Dev	251000	351.0	245.7	4.43
5th percentile	0.968 (*Tula Energy Distributing, 2011*)	0.05 (*Kirvoenergosbyt Pref., 2010*)	0.04 *(Smolenskenergosbyt pref., 2011)*	0.22 (*Magnitogorsky Iron Steel Works, 2008*)
1st quartile	20.77 (*Taganrog Metal Plant, 2011*)	0.50 (*Nizhny Novgorod Ret. Co., 2010*)	0.66 (*Nizhny Novgorod Ret. Co., 2008*)	0.62 (*Interregional Distribution Grid Co Centre*, 2008)
3rd quartile	1318 (*Sistema JSFC., 2007*)	31.07 (*Moscow Heating Network, 2010*)	23.25 (*Raspadskaya, 2008*)	2.8 (*Perm Energy Distributing Co., 2006*)
95th percentile	48257 (*Transneft Pref., 2008*)	306.91 (*Rushydro,2010*)	203.63 (*Tatneft, 2010*)	6.9 (*Lipetsk Energy Retail Co., 2007*)
# of obs.	1354	1135	699	834

Table 1. Descriptive statistics for the full period, 1998–2011 (continued)

	Percentage zero returns	**Spread**	**LOT**	**GEV**	**Amihud**
Mean	0.235	0.073	0.150	0.142	1.600
Median	0.109	0.026	0.041	0.014	0.530
St. Dev	0.274	0.174	0.298	0.352	2.125
5th percentile	0.004 *(Lukoil, 2001)*	0.0006 *(LUKOIL, 2001)*	0.003 *(VTB Bank, 2007)*	0.0001 *(Businessactive, 2009)*	0.00005 *(Rostelecom pref., 2007)*
1st quartile	0.026 *(JSC "OGK-1", 2007)*	0.008 (*Mostotrest, 2011*)	0.014 (*RAO Energy Sy. of East, 2011*)	0.002 (*Raspadskaya, 2007*)	0.014 (*Volga TGC, 2011*)
3rd quartile	0.371 *(Perm Energy Ret., 2010)*	0.075 (*Samarenergo, 2004*)	0.130 (*Tomsk En. Distg. Co., 2009*)	0.085 (*Astrakhan Energy Retail, 2008*)	2.595 (*Sredneuralsky Copsm. Pl., 2009*)
95th percentile	0.851 *(Ivenergosbyt, 2009)*	0.262 (*Saratovenergo, 2000*)	0.695 (*MC Strategia, 2009*)	0.849	6.107 (*Kolenergosbyt, 2009*)
# of obs.	1329	1339	1354	1352	1353

The empirical distribution of annual spreads seems to be strongly right-skewed: whereas the median of 2.6% is similar to the findings on developed and advanced developing markets (e.g. Lesmond (2005) reports 3% for Greece 1988–2000 and 2.4% for Malaysia 1987–2000), the mean is about three times higher, 7.3 percentage points, and almost coincides with the third quartile. While 95% of recorded annual spreads are below 26.2%, our dataset contains several observations at the cap level of 300%. Our findings indicate considerable liquidity improvement in Russia, as Lesmond (2005) documents an average spread of 47% for stocks, traded on Russian Trading System in 1994–2000.

LOT measure of transaction costs is on average twice larger than spread, but has a similar distribution pattern: the average of 15% is two percentage points higher than the third quartile. Here our findings are closer to those of Lesmond (2005) for 1994–2000, where he finds average LOT of 21 percentage points, however our median value of 4.1% is about three times smaller than reported by Lesmond (2005). Thereby in our data 95% of estimated LOT round-trip transaction costs is below 65% and still some very high outliers are present, hence LOT displays an even fatter right tail, than quoted spread.

Our GEV measure of the round-trip transaction costs is on average one percentage point lower than LOT and still strongly overshooting the quoted spread. However the median of our measure of 1.4% is three times smaller than that of LOT and even smaller than the median of quoted spreads. Our measure is very volatile, with standard deviation of 35 percentage points double as large as for quoted spreads and somewhat larger than of LOT. Right-skewness of GEV distribution is even more pronounced: whereas three fourths of estimated values lie below 8.5%, the reported mean is 14%. The most liquid company-year observation of the most illiquid quarter of our sample is a regional energy retail company (Astrakhan Energy Retail) in a crisis year (2008).

According to our estimates of price-impact measure of Amihud, a one-million-ruble buy order would lead to almost tripling the stock price on average and to a price increase of 53% of a median company. The second quarter of mostly "deep" stock-years starts with a regional electricity generating company in 2011, which price would change 1.4% as a reaction to one-million-ruble order volume. The impact of the same volume on the stock of Sredneuralsky Coppermelting Plant in 2009, which opens the quarter of the most "shallow" stock-years, would be almost quadrupling of the price (+260%). For some 67 least "deep" stock-years the price impact of one-million-ruble volume would be higher than seven times.

Median annual trading volume in our sample is 156 million rubles, or about 0.6 million rubles per day. Hence, for median stock the Amihud measure reflects price impact of 1.7 average daily volumes. Trading volume of 5% least traded stocks is however below 1 million per annum, or some 4 thousand rubles per day, so that Amihud measure is 250 (or more) times daily absolute return, what can easily lead to extremely high values. The upper quartile (or 338 stock-year observation with the highest trade value) starts with annual trading volume of 1.3 billion rubles, or about 5.4 million rubles per day. Thereby 5% most heavily-traded stock-years display trade volumes of 48 billion rubles (200 million rubles per day), as for Transneft preferred stock in 2008, and above.

5. Econometric Technique

In order to fully exploit our data we run panel regressions:

$$S_{it} = \alpha + \beta' X_{it} + \mu_i + \lambda_t + v_{it}, \qquad (10)$$

where μ_i denotes cross-sectional individual effects, λ_t denotes year effects and v_{it} is an idiosyncratic error term.

We rely on the standard technique in the asset pricing literature, the Fama-MacBeth (1973) regression, when analyzing the impact of transaction costs on the cross-sectional variation of returns. It is based on the assumption that expected returns of stocks are fully described by the linear combination of risk premia and factor loadings for all relevant factors:

$$E\left[Z_i\right] = \boldsymbol{\lambda}' B_i,$$

whereby $Z_{it} = r_{it} - r_{ft}$ denotes excess return, $\boldsymbol{\lambda}'$ is a transposed vector of risk-premia, and B_i is a vector of factor loadings or risk characteristics of company *i*. Given the values of factor loadings for each stock in each period the risk premia are estimated running *T* cross-section regressions (one for each period) and averaging the estimates:

$$Z_{it} = \boldsymbol{\lambda}_t' B_{it}$$

$$\overline{\boldsymbol{\lambda}} = \frac{1}{T}\sum_{t=1}^{T} \hat{\boldsymbol{\lambda}}_t$$

The corresponding standard errors for each *k*-th element of the risk-premia vector are calculated from the corrected time variance of the estimated premia:

$$\mathrm{var}\left[\lambda_{kt}\right] = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\lambda}_{kt} - \overline{\lambda}_k\right)^2$$

$$\mathrm{stderr}\left[\overline{\lambda}_k\right] = \sqrt{\mathrm{var}\left[\lambda_{kt}\right]\cdot\left(1 + \overline{f}'\,\Omega_f \overline{f}\right)/T},$$

where $\overline{f}$ denotes the vector of mean factor values and Ω_f is the factor variance-covariance matrix. For the risk factor *k* to be priced the corresponding risk premium should be significantly different from zero.

6. Results

6.1. Precision of indirect measures

We analyze the precision of our liquidity measure following Goyenko et al. (2009) using two type of indicators: average prediction error and correlation with benchmark. We use quoted closing spreads as benchmark. Results are reported in Table 2. Mean bias is not reported for Amihud measure and log of the trading volume, as these values have scales which differ from the one of quoted spread. Both LOT and GEV measures on average overestimate spreads, but GEV's bias is on average one percentage points lower as for the whole sample, as for both sub-samples as well. Admittedly, both measures of roundtrip transaction costs should be higher than quoted spreads, as they also include commissions, but the magnitude of reported bias is times higher than known commissions. The size of the bias for LOT measure we observe on Russian data is slightly higher than the reported by Goyenko et al. (2009, Table 3, Panels C–D) for US stocks: 0.07–0.08 vs. 0.05–0.06 respectively, which could be caused by higher idiosyncratic volatility of stock returns in our data. Interestingly, median bias (reported in table 2 in square brackets below the corresponding reported average error) is much lower: median prediction error for LOT is one percentage point, whereas for GEV only about 50 basis points.

In terms of cross-sectional correlation LOT and GEV seem to outperform volume based measures. LOT has higher correlation to the benchmark than GEV in the first sub-sample, but our measure outperforms LOT in 2008–2011, with reported correlation to spread of 0.72 to 0.69. The magnitude of coefficients for the whole sample, 0.47–0.49, coincides with values reported by Goyenko et al. (2009, Table 3, Panel A) for US data for LOT and their modification of it, 0.41–0.54. In 2008–2011 on MICEX zero return based indirect measures seem to proxy spreads better than for S&P500 stocks in 1993–2005.

To sum up, indirect liquidity measures seem to be reasonably precise for Russian data, where GEV is slighter more accurate than others, and can be used for analysis of liquidity, especially when data on benchmark is not available.

Table 2. Precision of indirect liquidity measures (average error / correlation) at spread

	LOT		**GEV**		**Amihud**		**Log(volume)**	
	Av. error	corr	Av. error	corr	Av. error	corr	Av. error	corr
1998–2011	0.07 [0.01]	0.49	0.06 [–0.004]	0.47		0.43		-0.43
1998–2007	0.08 [0.01]	0.48	0.07 [–0.001]	0.45		0.42		-0.43
2008–2011	0.07 [0.01]	0.69	0.06 [–0.005]	0.72		0.71		-0.68

6.2. Explaining liquidity

To test the hypothesis of positive impact of size, trading activity and negative impact of financial distress on liquidity we run a battery of panel regressions of the type $S_{it} = \alpha + \beta' X_{it} \left(+\mu_i + \lambda_t \right) + v_{it}$, where S_{it} is a liquidity measure of a stock i in year t (which is either GEV, LOT, Amihud, quoted spread or log volume), vector X contains values of explanatory variables for the corresponding observation. μ_i and λ_i denote possibly included individual cross-section and period effects. To ensure that explanatory variables are predetermined, we lag them one period. Market capitalization and trading volume, even in logs, are known to be non-stationary. However in our unbalanced panel set-up we have on average about 4 observations per stock, which are not necessarily subsequent, what reduces the importance of this issue. Moreover, all standard tests reject unit-root for both series. Therefore we treat log trading volume and log market capitalization as if they were stationary.

To exploit the full range of data we run panel regressions. For our round-trip transaction costs measure GEV standard F-test for redundant fixed effects allows using pooled data. However, pooled estimation yields statistically significant residual autocorrelation, even though not economically very large (about 10%). Therefore we report results of pooled least squares regression with heteroscedasticity adjusted weights and heteroscedasticity and autocorrelation adjusted White standard errors, as well as results of Arellano-Bond (1991) system general method of moments estimation. We provide estimates with and without period effects, as well as GMM estimates for the two sub-periods 1998–2007 and 2008–2011.

In all full-sample regressions (Table 3, Columns (1)–(6)) size, volume and market-to-book have the predicted sign, but market-to-book ratio turns out to be insignificant in all but two specification, where it is statistically significant on the 10% level, but economically negligible. An alternative measure of distress, previous year return, is insignificant throughout all full sample specifications, in contrast to results reported for an early auction market in Burhop and Gelman (2012). In pooled estimation set-up, previous year size and trading activity explain 15% of the GEV transaction cost measure variance (Table 3, column (1)), whereas period effects explain about 3% of liquidity variation (column (2)). According to the pooled specification with period effects (column (2)), which mitigates the consequences of capitalization growth with time, size and trading activity are highly significant and seem to have an impact of similar magnitude. Stocks of companies, which have a one unit higher log market value (which corresponds at the mean to 15.6 billion rubles higher market capitalization) have on average one percentage point lower transaction costs.

Table 3. Panel regression to explain liquidity (GEV as dependent)

	(1)PA	(2)PA	(3)PA	(4) GMM	(5) GMM	(6) GMM	(7)GMM 98-07	(8) GMM 08-11
Constant	32.4***	31.3***	31.2***	48.9***	49.0***	43.0***	45.6**	40.0***
	(2.4)	(2.4)	(2.4)	(0.20)	(7.8)	(8.1)	(21.9)	(8.8)
S_{it-1}				0.0	0.0	0.0	0.0	0.0
				(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
$Log(MC_{it-1})$	-1.0***	-1.0***	-1.0***	-1.8***	-1.8***	-1.6***	-0.9	-1.7***
	(0.3)	(0.3)	(0.3)	(0.5)	(0.5)	(0.5)	(0.7)	(0.6)
MTB_{it-1}	-0.2	-0.1	-0.2	-0.0*	-0.0*	-0.0	-1.5	-0.0
	(0.1)	(0.1)	(0.1)	(0.0)	(0.0)	(0.0)	(0.9)	(0.0)
$lnTV_{it-1}$	-0.8***	-0.8***	-0.8***	-1.2***	-1.2***	-1.0***	-1.4**	-0.8***
	(0.2)	(0.2)	(0.2)	(0.3)	(0.3)	(0.3)	(0.7)	(0.3)
R_{it-1}		-0.0	0.0		0.0	-0.0	0.8***	-0.0
		(0.4)	(0.4)		(0.0)	(0.0)	(0.1)	(0.0)
ρ_{it-1}			0.5			-1.8	1.6	-3.4
			(2.6)			(3.4)	(4.0)	(4.1)
Time effects	Y	Y	Y	Y	Y	Y	Y	Y
Firm effects	N	N	N	Y	Y	Y	Y	Y
Obs.	846	832	822	793	793	773	191	582
Stocks	211	211	211	207	207	207	83	202
R^2	0.18	0.18	0.19	-				
Hansen				0.98	0.98	0.97	2.46	0.96
				(0.32)	(0.32)	(0.32)	(0.12)	(0.33)

Estimates of LS pooled models as well as Arellano-Bond (1991) System GMM for the transaction costs (GEV measure, in percentage points) for the sample period from 1998 to 2011 of the type: $S_{it} = \alpha + \beta' X_{it} \left(+\mu_i + \lambda_t\right) + v_{it}$. White period standard errors are reported in parenthesis (except for Hansen statistic, where p-value is reported). Values marked with ***, ** and * are significant at 1%, 5% and 10% level respectively. R^2 is calculated as one minus the fraction of the residual variance to the variance of the dependent variable.

One standard deviation increase in log market cap (which is about 97 billion ruble increase in capitalization) reduces our illiquidity measure by 2.5 percentage points decrease. An increase of log volume by one unit, what corresponds at the mean to increasing annual trading volume by 500 million rubles, leads to 80 basis points decrease of round-trip transaction costs. One standard deviation increase of log volume, which is equivalent to raising annual trading volume by 9 billion, reduces transaction costs by 2.8 percentage points. Daily return autocorrelation – our proxy for insider trading – has a predicted sign, but is insignificant at all conventional levels.

Arrelano-Bond (1992) system GMM approach yields valid models: Hansen statistic does not allow to reject the validity of overidentifying restrictions in all cases (see columns (4)–(8)). As the approach involves differencing the dependent and explanatory variables, cross-section effects are automatically accounted for. GMM results for the full sample (columns (4)–(6)) reinforce our earlier findings: both size and trading activity stay statistically significant at the 1% significance level and of the same sign. Moreover, impact magnitude grows for size almost twice and 1.5 times for trading activity. However, these changes do not exceed two standard errors of the corresponding estimates. Lagged transaction costs are statistically and economically insignificant.

We break our sample into two sub-periods: 1998–2007 and 2008–2011. Even though the first sub-period is longer, it contains substantially less observations, due to a smaller number of traded stocks, as well as less observed years of trading per stock. Whereas the results for the second sub-period are qualitatively the same as for the whole period, in the first sub-period the proxy of financial distress – previous year return – gains statistical significance, even though the sign is opposite to the hypothetical, while market capitalization becomes insignificant. Insider trading seems to have no impact on transaction costs in both sub-periods. However, the magnitude of all coefficients grows in the first sub-period and the results seem to be less stable.

Results for the traditional LOT measure in Table 4 are qualitatively very similar: market capitalization and turnover value significantly reduce transaction costs the following year. This finding is significant on 1% level for the full sample and second sub-period in all model modifications. In the first sub-period, while market-to-book ratio becomes significant, the size proxy however still remains significant at the 10% level. The magnitude of coefficients is comparable to GEV model estimates. The major difference is the growing explanatory power of the regression, reaching 25% in pooled estimation with period effects. This may be due to smaller skewness of LOT measure compared to GEV. Informed trading does not influence liquidity significantly, except for the 2008–2011 sub-period, where it enhances liquidity on a 10% significance level, which is a rather puzzling result. Possibly our proxy for informed trading in this sub-sample catches primarily the spread size over negative autocorrelation, as outlined in Roll (1984), and therefore conveys less information about the impact of insiders.

Using Amihud measure we can reinforce our findings. Except for the magnitude of coefficients and minor changes in sub-period result estimates' significance, regressions reveal supporting evidence of liquidity enhancing properties of size and trading activity. A liquidity-enhancing effect of daily return autocorrelation is even stronger (and in three specifications significant at 5% level or higher), than in LOT regressions, which indicates our proxy capturing rather bid-ask-bounce than informed trading. *R*-squared of the pooled regression with period effect doubles to almost reach 50%.

Summarizing our study of liquidity drivers we can conclude that size and previous period trading activity play the major role, whereas financial distress or amount of informed trading seems to have no influence on liquidity. This result supports information asymmetry theory to some extent, as public information is more abundant for larger companies and information discovery occurs also through trading, but does not support our theory of negative impact of informed trading, possibly due to a weak proxy.

Table 4. Panel regression to explain liquidity (LOT as dependent)

	(1)PA	(2)PA	(3)PA	(4) GMM	(5)GMM	(6) GMM	(7)GMM 98-07	(8) GMM 08-11
Constant	36.8***	35.7***	33.8***	43.6***	46.4***	44.2***	40.4***	40.1***
	(2.2)	(2.1)	(2.0)	(5.1)	(7.3)	(6.3)	(10.0)	(6.4)
S_{it-1}				0.0**	0.0**	0.0	0.0*	0.0
				(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
$Log(MC_{it-1})$	-0.8***	-0.8***	-1.0***	-1.4***	-1.6***	-1.7***	-0.6	-1.8***
	(0.3)	(0.3)	(0.3)	(0.4)	(0.5)	(0.5)	(0.5)	(0.6)
MTB_{it-1}	-0.0	-0.0	-0.0	-0.0	-0.0	0.0	-1.0**	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.5)	(0.0)
$lnTV_{it-1}$	-1.1***	-1.1***	-0.9***	-1.2***	-1.2***	-1.0***	-1.3***	-0.7***
	(0.2)	(0.2)	(0.1)	(0.2)	(0.3)	(0.2)	(0.5)	(0.2)
R_{it-1}		-0.2	-0.1		-0.0	-0.0	0.8***	-0.0
		(0.3)	(0.3)		(0.0)	(0.0)	(0.1)	(0.0)
ρ_{it-1}			0.3			-3.4	2.3	-5.7*
			(2.2)			(2.9)	(3.2)	(3.3)
Time effects	Y	Y		Y	Y	Y	Y	Y
Firm effects	N	N		Y	Y	Y	Y	Y
Observations	884	870	961	883	827	818	197	618
Stocks	212	212	212	209	208	208	84	205
R^2	0.25	0.25	0.29	-				
Hansen				0.01	0.66	0.94	5.47	2.09
				(0.93)	(0.42)	(0.33)	(0.07)	(0.15)

Estimates of LS pooled models as well as Arellano-Bond System GMM for the transaction costs (LOT measure, in percentage points) for the sample period from 1998 to 2011 of the type: $S_{it} = \alpha + \beta' X_{it} \left(+ \mu_i + \lambda_t \right) + v_{it}$. White period standard errors are reported in parenthesis (except for Hansen statistic, where p-value is reported). Values marked with ***, ** and * are significant at 1%, 5% and 10% level respectively. R^2 is calculated as one minus the fraction of the residual variance to the variance of the dependent variable.

Table 5. Panel regression to explain liquidity (Amihud as dependent)

	(1)PA	(2)PA	(3)PA	(4) GMM	(5)GMM	(6) GMM	(7)GMM 98-07	(8) GMM 08-11
Constant	490.7***	486.0***	460.9***	722.0***	742.1***	689.0***	-27.4	838.1***
	(23.4)	(22.5)	(23.0)	(149.7)	(169.0)	(160.6)	(94.9)	(151.3)
S_{it-1}				-0.12	-0.17	-0.15	0.42**	-0.21*
				(0.15)	(0.16)	(0.16)	(0.20)	(0.12)
S_{it-2}				-0.16***	-0.15***	-0.15***		-0.18***
				(0.04)	(0.04)	(0.04)		(0.05)
$Log(MC_{it-1})$	-21.0***	-19.8***	-20.6***	-28.9***	-30.0***	-30.1***	-17.4***	-32.3***
	(2.7)	(2.6)	(2.6)	(5.2)	(5.4)	(5.3)	(5.5)	(6.4)
MTB_{it-1}	-0.0	-0.0	-0.0	8.1	8.3*	8.2*	-4.1	9.2*
	(0.0)	(0.0)	(0.0)	(5.1)	(4.9)	(5.0)	(3.4)	(5.0)
$lnTV_{it-1}$	-10.3***	-10.5***	-9.1***	-17.8***	-18.0**	-15.6**	8.1	-20.0***
	(1.7)	(1.6)	(1.7)	(6.4)	(7.1)	(6.8)	(9.4)	(6.8)
R_{it-1}		-6.4*	-4.1		0.1	0.1	2.7***	0.1***
		(3.5)	(3.5)		(0.0)	(0.0)	(0.9)	(0.0)
ρ_{it-1}			-90.7***			-85.2**	-77.2	-90.7**
			(21.7)			(37.3)	(51.7)	(41.1)
Time effects	Y	Y	Y	Y	Y	Y	Y	Y
Firm effects	N	N	N	Y	Y	Y	Y	Y
Observations	899	885	885	710	701	701	218	554
Stocks	211	211	211	193	193	193	86	190
R^2	0.48	0.48	0.49	-				
Hansen				4.46	4.25	3.96	2.27	1.06
				(0.11)	(0.12)	(0.14)	(0.13)	(0.59)

Estimates of LS pooled models as well as Arellano-Bond System GMM for the illiquidity measure (Amihud measure, in percentage points) for the sample period from 1998 to 2011 of the type: $S_{it} = \alpha + \beta' X_{it} \left(+\mu_i + \lambda_t \right) + v_{it}$. White period standard errors are reported in parenthesis (except for Hansen statistic, where p-value is reported). Values marked with ***, ** and * are significant at 1%, 5% and 10% level respectively. R^2 is calculated as one minus the fraction of the residual variance to the variance of the dependent variable.

6.2. Liquidity and asset prices

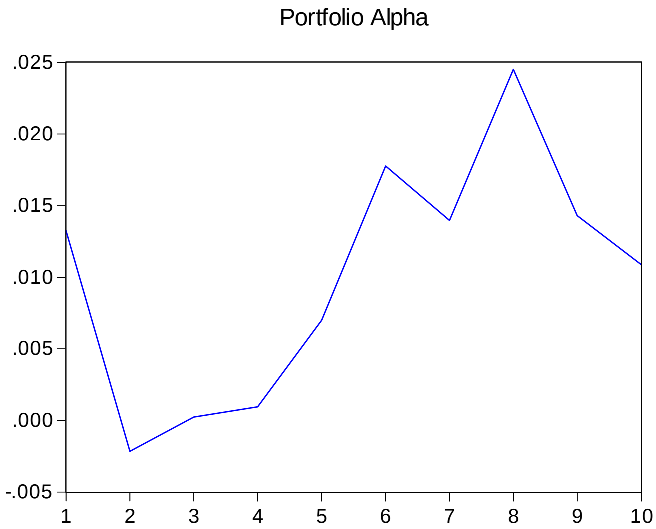
To analyze the impact of liquidity on asset pricing on the Russian stock market we perform two types of tests: first, we use liquidity sorted portfolios and analyze alphas of Carhart (1997) four-factor model to show that they positively depend on illiquidity; second, we utilize the full set of individual security returns and use illiquidity measures as characteristics to find whether they yield a significant premium.

Table 6. Results of portfolio-based time-series asset pricing regressions

	P1	P3	P5	P8	P10	P10-P1	L30-U30
Alpha	.0132**	.0002	.0070	.0245**	.0109	-.0091	.0125**
	(.0054)	(.0064)	(.0081)	(.0117)	(.0136)	(.0148)	(.0060)
Market β	1.15***	0.93***	0.84***	1.00***	0.45**	-0.69***	-0.34***
	(0.10)	(0.11)	(0.11)	(0.23)	(0.20)	(0.24)	(0.12)
SMB β	0.06	-0.11*	0.10	0.67***	0.67***	0.63***	0.56***
	(0.07)	(0.07)	(0.12)	(0.24)	(0.19)	(0.22)	(0.10)
HML β	-0.05	0.06	0.15	-0.09	-0.27	-0.23	-0.15
	(0.07)	(0.09)	(0.10)	(0.20)	(0.18)	(0.19)	(0.11)
Mom. β	.01	-0.19**	0.04	0.05	-0.40*	-0.40**	-0.27*
	(0.06)	(.08)	(0.11)	(0.20)	(0.21)	(0.20)	(0.16)
R^2	0.77	0.60	0.38	0.25	0.29	0.45	0.56
Av. # of stocks	11	11	11	11	11		
# of obs. T	138	138	138	138	138	138	138

Estimates of the Carhart (1997) time-series regressions for the sample period from 2000 to 2011 for a selection of one-way sorted decile liquidity portfolios (from most liquid (P1) to least liquid (P10)), for a combination of a long position in the least liquid decile and a short position in the most liquid decile, as well as for a combination of a long position in the least liquid 30% of stocks and a short position in the 30% most liquid stocks. Newey-West heteroscedasticity and autocorrelation adjusted standard errors are reported in parenthesis. Values marked with***, ** and * are significant at the 1%, 5%, and 10% level respectively.

Figure 12. Alpha values of liquidity-sorted portfolios



In the first step we analyze returns of the ten liquidity-sorted portfolios using the standard time-series approach of testing the Carhart (1997) model. Therefore we run time-series regressions (one for each portfolio) of portfolio returns on the Carhart's four factors: market return and returns of "small minus big" (SMB), "high book-to-market minus low book-to-market" (HML) and momentum factor mimicking portfolios. Selection of results is presented in Table 6. Market betas are significant for all analyzed portfolios (reported in columns (1) through (5) as well as unreported), and is statistically indistinguishable from unity for all portfolios but the least liquid decile (column 5). In several cases portfolios have significant positive loadings on SMB factor (e.g. column (4) and (5)), and for portfolio 3 we find negative beta on momentum, which is significant on the 5% level. However, the focus of this test is on the regression intercepts, which reflect the average pricing errors of the model. Beside reporting selected values of alphas in the first row of table 6, we report all of them in figure 12. Gibbons-Ross-Shanken (1989) test of joint equality of pricing errors to zero yields a statistic of 2.1, which allows us to reject the null of no pricing errors at 5% significance level (p-value = 0.025).[1] Looking at graph 1 one can suggest that alphas grow with illiquidity, however the relationship seems to be non-linear, due to the high alpha of the most liquid portfolio and low alpha of portfolios P9 and P10. Therefore going long in the least liquid decile and short in the 10% most liquid stocks does not provide a significant Carhart's alpha (see column (6)). However, extending the strategy to 30% of stocks with the highest and 30% with the lowest round-trip transaction cost estimates yields a pricing error of 1.25% per month, which is significant at the 5% significance level (see column 7). Such portfolio exhibits significant factor loadings: a negative one-third on market risk and a positive loading

[1] Details are available on demand.

of more than one half on the size factor, both significantly different from zero on the 1% level. On the 10% significance level is the negative sensitivity of the portfolio to momentum factor. Still, accounting for all these risks, the strategy earns 15% annually, before transaction costs. Transaction costs would however eliminate profits in this case: average round-trip transaction costs of the 30% least liquid stocks in 2010 were, according to GEV measure, about 18%. Somewhat more sophisticated strategies, which allow avoiding very illiquid stocks, would yield significant alphas not compensated by transaction costs. E.g. a long position in P6 and a short position in P2 yield an alpha of about 16% p.a., whereas corresponding transaction costs were on average about 9.6% during this period.[2] A net alpha of 6% p.a. suggests a premium not only for the level of liquidity, but for the liquidity risk, too.

To test the impact of liquidity as risk factor on asset pricing we run a GMM asset pricing regression in the discount-factor form:

$$m_{t+1} R^e_{i,t+1} = 0,$$

where $R^e_{i,t+1}$ denotes excess return of portfolio i at $t+1$ and $m_{t+1} = \left(1 - \boldsymbol{b}'\boldsymbol{f}_{t+1}\right)$ being a linear stochastic discount factor. $\boldsymbol{f}$ denotes a vector of risk-factor realizations and $\boldsymbol{b}$ is a vector of penalty parameters. We form the liquidity risk factor as excess return on a portfolio long in 30% least liquid stocks and short in 30% most liquid stocks. We test its impact along with market risk and compare the result to including the remaining three Carhart's factors.

Table 7. Results of GMM stochastic discount factor regressions

	(1)	**(2)**
Market *b*	3.08***	5.79***
	(1.07)	(1.66)
Liquidity *b*	2.13***	2.17
	(0.51)	(2.27)
SMB *b*		3.27
		(2.15)
HML *b*		2.77
		(2.62)
Mom. *b*		2.95
		(2.65)
Hansen1	7.35	1.76
	(0.50)	(0.88)
Hansen2		5.58
		(0.13)

Estimates of GMM regressions of the type $R^e_{i,t+1}\left(1 - b' f_{t+1}\right) = 0$ for the sample period from 2000 to 2011 for ten one-way sorted decile liquidity portfolios (from most liquid (P1) to least liquid (P10)). Values marked with ***, ** and * are significant at the 1%, 5%, and 10% level respectively.

[2] Details are available on demand.

In a parsimonious specification (Table 7, column 1) both coefficients are highly significant, indicating that both factors are priced. The reported penalty criteria according to $\Omega_f b = \lambda$, where Ω_f is the factors' variance-covariance matrix, correspond to risk-premia of 1.4% and 1.5% respectively (or 17% and 18% p.a.). Hansen statistic of 7.35 with 8 overidentifying restrictions clearly indicates that the model is valid. The most general specification (column 2) appears to have valid restrictions either (p-value of the Hansen statistic of 0.88), but only the market risk factor has a significant penalty parameter. Direct comparison of the two models (*Hansen2* statistic) does not allow rejecting the null of equivalent fit; therefore the first model is preferable. Hence, market risk and liquidity risk seems to be priced on the Russian stock market.

Admittedly, the above test design, primarily the portfolio formation, favors acceptance of liquidity risk pricing. To reinforce this result on a broader basis we test for the impact of liquidity on Russian asset pricing using the full dataset of individual stock returns, treating liquidity as characteristic.

Table 8. Results of cross-sectional asset pricing regressions

	(1)	**(2)**	**(3)** **2008–2011**	**(4)**
Constant	-.0024	.0002	-.0044	-.0092
	(.0061)	(.0107)	(.0066)	(.0111)
Market beta $\bar{\lambda}_{\beta}$	.0184*	.0172	-.0031	.0259**
	(.0110)	(.0129)	(.0158)	(.0124)
SMB $\bar{\lambda}_{S}$	-.0012	.0082	.0033	
	(.0150)	(.0193)	(.0128)	
HML $\bar{\lambda}_{HML}$	.0190	.0268	.0035	
	(.0150)	(.0184)	(.0120)	
Momentum	.0038	.0001	-.0010	
	(0.150)	(.0190)	(.0175)	
Illiquidity (GEV)		-.0246	.0358*	.0200
TC		(.0331)	(.0215)	(.0192)
Average R^2	0.20	0.28	0.13	0.10
Average # of stocks	120	99	216	99
# of cross-sections *T*	144	144	48	144

Estimates of the Fama-MacBeth (1973) regressions for the sample period from 2000 to 2011. Reported coefficient values $\bar{\lambda}_k$ are averages of 144 (48 for column (3)) regression estimates of the type: $Z_{it} = \alpha_t + \lambda_t' B_{it} + u_i$, where λ_t' denotes the transposed vector of risk premia and B_{it} denotes the vector of risk factor loadings, which serve as explanatory variables in each cross section. Standard errors are calculated as, according to the Fama-MacBeth (1973) procedure with the Shanken (1992) correction (see Section V), and are reported in parentheses. Values marked with ***, ** and * are significant at the 1%, 5%, and 10% level respectively. Average R^2 is an arithmetic mean of R^2 for each cross-section.

Results of cross-sectional asset-pricing regressions are summarized in Table 8. Out of the four Carhart factors only market risk seems to matter, as it yields a statistically significant premium at least on the 10% level in Carhart's (1997) model as well as in a liquidity-augmented version of CAPM (Table 8, columns 1 and 4 respectively). Monthly market risk premium is about 1.8–2.6% per month or about 21.6–1.2% on annual basis, which is more than double of the values reported for developed markets. Thereby size, book-to-market and momentum factors yield no statistically significant premia. Transaction costs yield a marginally significant premium (on the 10% level) only in the second sub-sample (column 3), of about 3.6% on the monthly basis or about 43% per annum, which is much less than found by Burhop and Gelman (2012) for the Berlin Stock Exchange in 1892–1913 and obtained by Acharya and Pedersen (2005) for the US value-weighted portfolios in 1964–1999. The size of the premium implies that net returns across companies are equal if stocks are traded once in about 2.5 years. The illiquidity measure has a smaller magnitude, than sensitivity to market risk: it is on average 0.14 (see Table 1) compared to average beta of 0.78. That means that an average company's risk premium for illiquidity is about 6% p.a. and risk premium for market risk is about 17–23% p. a.

Using the full sample we fail to find supportive evidence of the illiquidity characteristic being priced, also in the parsimoniously specified liquidity augmented CAPM specification.

Hence, liquidity seems to be a priced factor on a Russian stock market, but not necessarily priced as characteristic.

7. Conclusion

We find that a wide-spread measure of transaction costs – Lesmond et al. (1999, LOT) – can be substantially improved to account for substantial idiosyncratic shocks and develop a corresponding measure (GEV), which significantly outperforms LOT in simulations.

We show that on the Russian stock market indirect measures of liquidity are sensible proxies of quoted spread, whereas the GEV measure developed in this paper is slightly preferable to the existing ones.

Using these measures we provide evidence of size and trading activity positively influencing liquidity on the Russian stock market. On the other hand, our results show that financial distress and informed trading proxies have no significant impact on liquidity. Thus the information asymmetry theory of liquidity is only partly supported.

Furthermore, liquidity turns out to be a statistically and economically significant factor in asset pricing on the Russian market. However, we do not find conclusive evidence of liquidity being priced as characteristic.

Our research can be extended in several ways. First, one could include other proxies of insider trading, such as mentioning of such in financial press or similar. Second, one could use our analysis of drivers of liquidity set-up to explore possible effects of cross-listing, as well as effects of existence of hedging instruments (i.e. futures and options). Both parameters are subject to policy decisions, which could lead to enhancement of liquidity for certain stocks or for the market as a whole.

References

Acharya, V.V., and Pedersen, L.H. (2005) Asset pricing with liquidity risk, Journal of Financial Economics, 77, 375–410.

Amihud, Y. (2002) Illiquidity and stock returns: cross-section and time-series effects, Journal of Financial Markets, 5, 31–56.

Arellano, M. and Bond, S., (1991) Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, Review of Economic Studies, 58, 277–97.

Bekaert, G., Harvey, C., and Lundblad, C. (2007) Liquidity and Expected Returns: Lessons from Emerging Markets. Review of Financial Studies, 20, 1783–831.

Burhop, C., and Gelman, S. (2011) Liquidity Measures, Liquidity Drivers and Expected Returns on an Early Call Auction Market, MPI Collective Goods Preprint, No. 2011/19.

Campbell, J.W., Lo, A.W., and MacKinlay, A.C. (1997) The Econometrics of Financial Markets, University of California Press.

Carhart, M. (1997) On Persistence in Mutual Fund Performance, The Journal of Finance, 52, 57–82.

Chordia, T., Huh, S.-W., and A. Subrahmanyam (2009) Theory-based illiquidity and asset pricing, Review of Financial Studies, 22, 3629–3668.

Chordia, T., Roll, R. and A. Subrahmanyam (2002) Order imbalance, liquidity and market returns, Journal of Financial Economics, 65, 111–130.

Eleswarapu, V.R. (1997), Cost of transacting and expected returns in the NASDAQ market, Journal of Finance, 52, 2113–2127.

Eleswarapu, V., and M. Reinganum (1993) The seasonal behaviour of the liquidity premium in asset pricing, Journal of Financial Economics, 34, 373–86.

Fama E.F., J. MacBeth (1973) Risk, return, and equilibrium: Empirical tests, Journal of Political Economy, 91, p. 607–636.

Gibbons, M.R., Ross, S.A. & Shanken, J., (1989) A test of the efficiency of a given portfolio, Econometrica, 57, 1121–1152.

Goyenko, R., Holden, C., and Trzcinka, C. (2009) Do liquidity measures measure liquidity? Journal of Financial Economics, 92, 153–181.

Kyle, A. S. (1985) Continuous Auctions and Insider Trading, Econometrica, 53, 1315–1335.

Lesmond, D. (2005) Liquidity of Emerging markets, Journal of Financial Economics, 77, 411–452.

Lesmond, D., Ogden, J., and Trzcinka, C. (1999) A new estimate of transaction costs, Review of Financial Studies, 12, 1113–1141.

Llorente, G., Michaely, R., Saar, G., and Wang, J. (2002) Dynamic Volume-Return Relation of Individual Stocks, Review of Financial Studies, 15, 1005–1047.

Pagano, M. (1989) Endogenous market thinness and stock price volatility, Review of Economic Studies, 56, 269–288.

Pastor, L. and Stambaugh, R.F. (2003) Liquidity risk and expected stock returns, Journal of Political Economy, 111, 642–685.

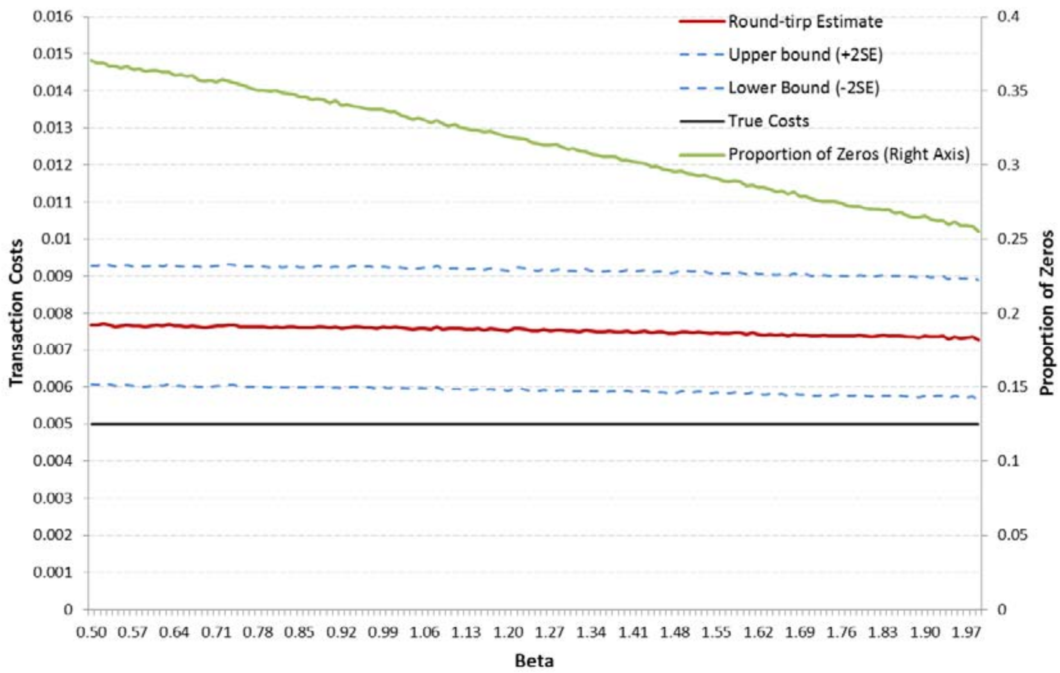
Rochet, J.-C., and Vila J.-L. (1994) Insider Trading without Normality, Review of Economic Studies, 61, 131–152.

Roll, R. (1984) A simple implicit measure of the bid-ask spread in an efficient market, Journal of Finance, 39, 1117–1140.

Shanken, J. (1992) On the Estimation of Beta-Pricing Models, Review of Financial Studies, 5(1), 1–33.

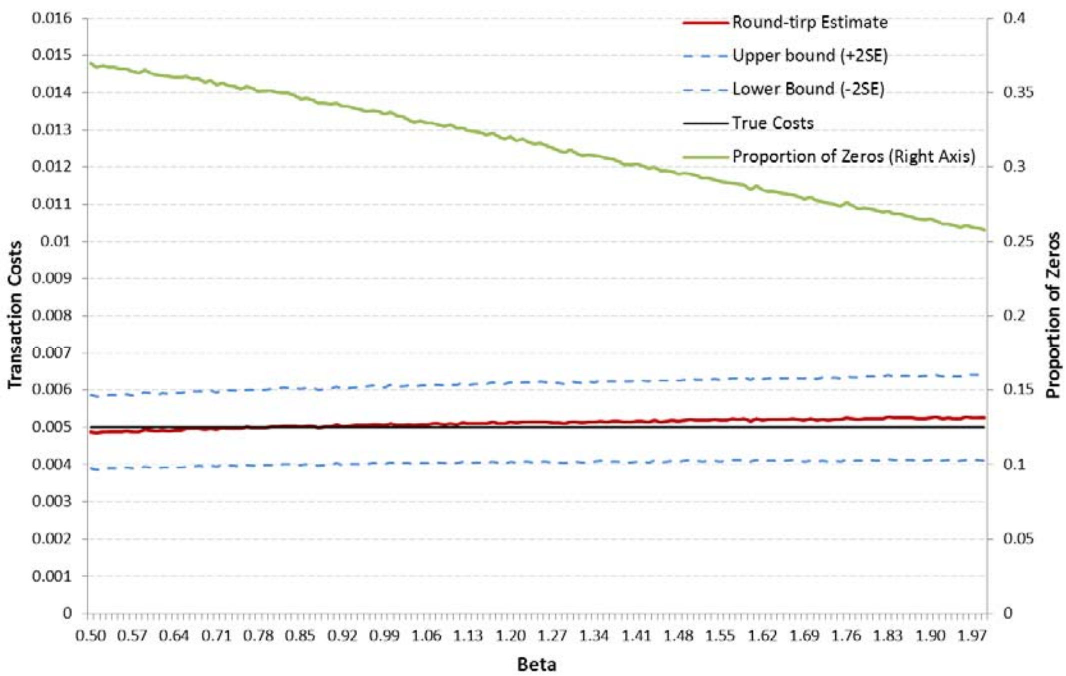
Appendix

Figure A1. The LOT Estimates and Beta



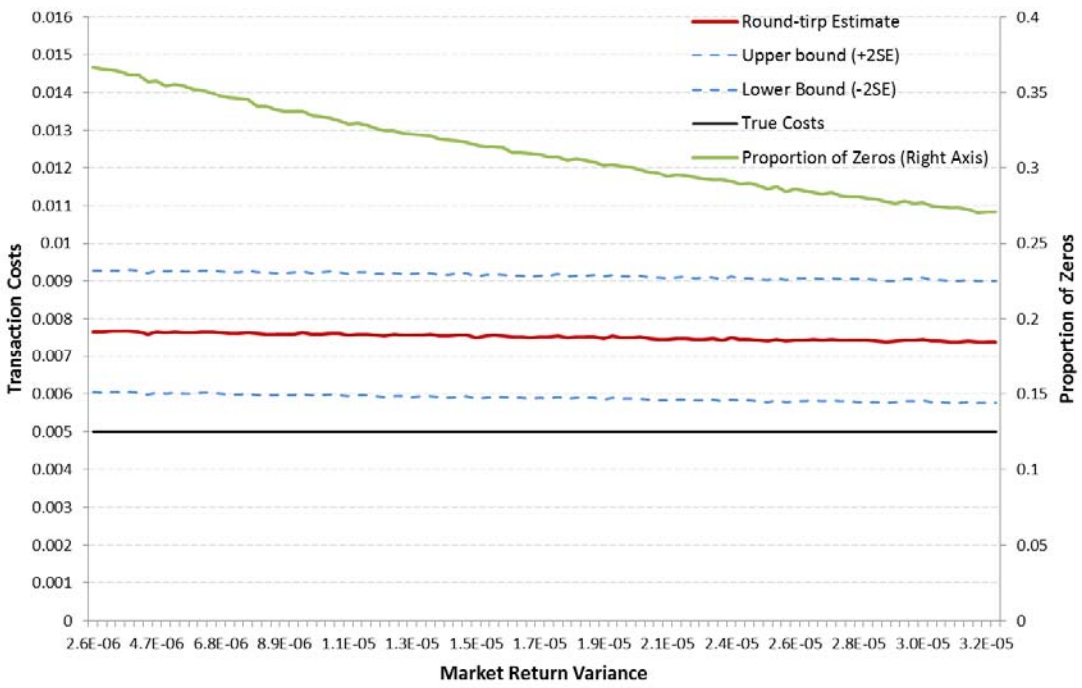
Transaction cost threshold $\alpha_2 - \alpha_1$ is set to 0.005, β ranges from 0.5 to 1.99 with increment of 0.01. The simulation structure is similar to one underlined in previous figures.

Figure A2. The GEV Estimates and Beta



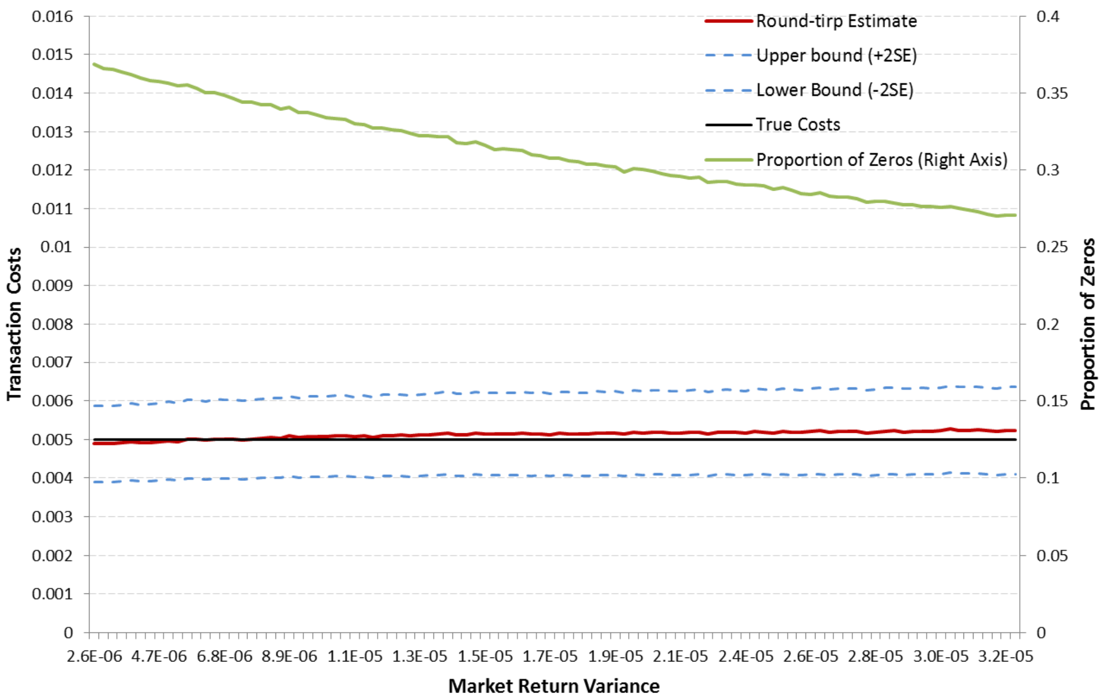
Transaction cost threshold $\alpha_2 - \alpha_1$ is set at 0.005, β ranges from 0.5 to 1.99 with increment of 0.01. The simulation structure is similar to one underlined in previous figures.

Figure A3. The LOT Estimates and Market Variance



The green solid line represents the variance of market return innovation term (scaled on the right axis). The market return variance increases each simulation step from 2.57×10^{-6} to 3.23×10^{-5} with 3×10^{-7} increment.

Figure A4. The GEV Estimates and Market Variance



The green solid line represents the variance of market return innovation term (scaled on the right axis). The market return variance gradually increases from 2.57×10^{-6} to 3.23×10^{-5} with 3×10^{-7} increment.

Appendix A5. Results of cross-sectional asset pricing regressions

	(1)	(2)	(3)	(4)	(5)
			08–12		
Constant	-.0024	.0031	-.0058	-.0118	
	(.0061)	(.0160)	(.0067)	(.0107)	
Market beta $\overline{\lambda}_{\beta}$	.0184*	.0131	-.0024	.0284**	
	(.0110)	(.0206)	(.0157)	(.0124)	
SMB $\overline{\lambda}_{S}$	-.0012	.0127	.0023		
	(.0150)	(.0310)	(.0128)		
HML $\overline{\lambda}_{HML}$	.0190	.0264	.0024		
	(.0150)	(.0290)	(.0121)		
Momentum	.0038	.0003	.0017		
	(0.150)	(.0313)	(.0176)		
Illiquidity (LOT)		-.0104	.0351	.0239	
TC		(.0546)	(.0236)	(.0214)	
Average R^2	0.20	0.25	0.13	0.09	
Average # of stocks	120	103	223	103	
# of cross-sections *T*	144	144	48	144	

Estimates of the Fama-MacBeth (1973) regressions for the sample period from 2000 to 2011. Reported coefficient values $\overline{\lambda}_k$ are averages of 144 (48 for column (3)) regression estimates of the type: $Z_{it} = \alpha_t + \lambda_t' B_{it} + u_i$, where λ_t' denotes the transposed vector of risk premia and B_{it} denotes the vector of risk factor loadings, which serve as explanatory variables in each cross section. Standard errors are calculated as, according to the Fama-MacBeth (1973) procedure with the Shanken (1992) correction (see Section V), and are reported in parentheses. Values marked with ***, ** and * are significant at the 1%, 5%, and 10% level respectively. Average R^2 is an arithmetic mean of R^2 for each cross-section.

Борисенко, Д. С., Гельман, С. В. Ликвидность, асимметрия информации и ценообразование активов на российском рынке акций : препринт WP9/2012/01 [Текст] / Д. С. Борисенко, С. В. Гельман ; Нац. исслед. ун-т «Высшая школа экономики». – М. : Изд. дом Высшей школы экономики, 2012 (на англ. яз.).

В статье исследуется влияние асимметрии информации и инсайдерской торговли на ликвидность и влияние ликвидности на ценообразование активов на российском фондовом рынке за период 1998–2011 гг. Используется набор существующих замещающих переменных для ликвидности, а также авторская модификация меры, предложенной Lesmond et al. (1999). Также используется положительный коэффициент автокорреляции дневных доходностей как индикатор инсайдерской торговли. Полученные результаты свидетельствуют о том, что асимметрия информации ухудшает ликвидность, однако не подтверждают гипотезу негативного воздействия информированной торговли, что может быть отчасти вызвано неоптимальным индикатором инсайдерской активности. Кроме того, ликвидность как и рыночные риски, является основным фактором ценообразования активов на российском фондовом рынке. Однако этот результат не является устойчивым к рассмотрению ликвидности как характеристики, а не как фактора.

Ключевые слова: трансакционные издержки; премия за ликвидность; информированная торговля.

Борисенко Дмитрий Сергеевич
Международный институт экономики и финансов
Национальный исследовательский университет – Высшая школа экономики
Покровский бульвар 11, 109028 Москва, Россия

Гельман Сергей Викторович
Международный институт экономики и финансов
Национальный исследовательский университет – Высшая школа экономики
Покровский бульвар 11, 109028 Москва, Россия
Email: sgelman@hse.ru

Препринт WP9/2012/01
Серия WP9
Исследования по экономике и финансам

Д.С. Борисенко, С.В. Гельман

Ликвидность, асимметрия информации
и ценообразование активов на российском рынке акций

(*на английском языке*)