Idiosyncratic risk and indirect transaction costs: enhancement of the Lesmond, Ogden and Trzcinka (1999) measure

Dmitry Borisenko

International College of Economics and Finance

National Research University Higher School of Economics

Ul. Shabolovka 26, 119049 Moscow, Russia

Email: dmitrijborisenko@yandex.ru

Sergey Gelman[†]

International College of Economics and Finance

National Research University Higher School of Economics

Ul. Shabolovka 26, 119049 Moscow, Russia

Email: sgelman@hse.ru

Abstract

A popular measure of liquidity – the indirect measure of full transaction costs introduced in Lesmond et al. (1999) – is known to be significantly upward biased. We show that the bias stems from the mistreatment of large countercyclical idiosyncratic shocks in individual stock returns. We suggest a modification, which applies a Generalized Extreme Value distribution to capture such shocks, and, with help of Monte-Carlo simulations, find that our approach fully eliminates bias under a normal distribution assumption and significantly diminishes bias under fatter tails. Using a sample of S&P 500 constituent stocks we show that our modified estimate is much more precise than the Lesmond et al. measure (1999).

[†] Corresponding author. Phone: +7 495 772 9590 ext. 026095, fax: +7 495 580-89-19

I. Introduction

Being able to correctly assess liquidity of a financial market is extremely valuable from both theoretical and practical points of view, as liquidity is a crucial factor in asset pricing, the most important objective in market microstructure design, and a vital parameter in developing trading strategies. Nevertheless, measuring liquidity of a financial market is a challenging task, especially if intraday data is unavailable. But being able to utilize only daily observations provides for liquidity analysis of a plethora of international stock markets with tremendous historical depth. Goyenko et al. (2009) show that a number of indirect measures of transaction costs and price impact based on daily data turn out to be good proxies of true liquidity. One of them – Lesmond, Ogden and Trzcinka (1999, further LOT), a measure of full transaction costs – requires only closing prices for stocks and for the market index under analysis. It exploits the idea that informed traders would act upon new information only if its value exceeded costs, incurred by trading on the information. Goyenko et al. (2009) show that this measure provides a reasonable correlation with benchmarks (intra-day measured effective and realized spreads), but considerably overshoots the benchmarks - by more than 4 percentage points - which is beyond the scope of fees, commissions and feasible price impact.

We argue that this bias is due to the incorrect treatment of large countercyclical idiosyncratic shocks. LOT presents returns as latent returns, fully revealing changes in the stock value, minus transaction costs. In this context the latent return is a combination of a linear function of the market return and a company-specific shock. The LOT estimation approach utilizes only observed market returns; large idiosyncratic shocks of an opposite sign are misinterpreted as very large transaction costs. We show that this flaw can be corrected by a simple modification of the measure which would treat such shocks separately, applying a Generalized Extreme Value distribution. We introduce our modification and show in a simulation study that our modified measure does not have bias under ideal assumptions for LOT, whereas LOT is significantly biased. If we simulate fat-tailed latent returns our measure becomes biased, but considerably less than the LOT measure. Finally, empirical comparison on S&P 500 constituent stocks provides for superiority of our measure over LOT as a proxy for spreads. Hence, we suggest using our modification instead of LOT in studies involving liquidity measures, estimated based on daily data. The rest of the paper is structured as follows: Section 2 discusses the LOT measure and its bias. Section 3 introduces our modification of the measure and discusses the simulation results. Section 4 gives an overview of the empirical study, followed by the conclusion.

II. LOT measure and its bias

An informed investor will trade only if the value of information 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:

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

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:

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

The corresponding log-likelihood is:

$$\begin{aligned} &LogL(\alpha_{1j}, \alpha_{2j}, \beta_j, \sigma_j \mid R_{jt}, R_{mt}) = \sum_{\Omega_0} ln \left[\Phi_{2j}\left(\frac{\alpha_{2j} - \beta_j R_{mt}}{\sigma_j} \right) - \Phi_{1j}\left(\frac{\alpha_{1j} - \beta_j R_{mt}}{\sigma_j} \right) \right] + \\ &+ \sum_{\Omega_1} ln \frac{1}{\left(2\pi\sigma_j^2\right)^{\frac{1}{2}}} - \sum_{\Omega_1} \frac{1}{2\sigma_j^2} \left(R_{jt} + \alpha_{1j} - \beta_j R_{mt}\right)^2 + \sum_{\Omega_2} ln \frac{1}{\left(2\pi\sigma_j^2\right)^{\frac{1}{2}}} - \sum_{\Omega_2} \frac{1}{2\sigma_j^2} \left(R_{jt} + \alpha_{2j} - \beta_j R_{mt}\right)^2 \end{aligned} \tag{3}$$

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.

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 individual stock returns on a growing market:

$$R_{jt} = \beta_j R_{mt} + \epsilon_{jt} - \alpha_{1j} < 0, \qquad R_{mt} > 0, \beta_j > 0.$$

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} in absolute value, that is

$$\epsilon_{jt} < -\beta_j R_{mt} + \alpha_{1j}, \quad R_{mt} > 0, \ \beta_j > 0.$$

On the contrary, to minimize the error for such observation taken alone LOT approach would set the buy-transaction cost on an unjustifiably high level:

$$\hat{\alpha}_{2j} = \alpha_{1j} - \epsilon_{jt} = \beta_j R_{mt} + x_{jt}, \ \text{ where } x_{jt} = -\left(\epsilon_{jt} + \beta_j R_{mt} - \alpha_{1j}\right) > 0.$$

Thus such observations distort the correct estimation of buy costs pushing the estimate upwards. The same intuition is applied for positive stock returns on a falling market.

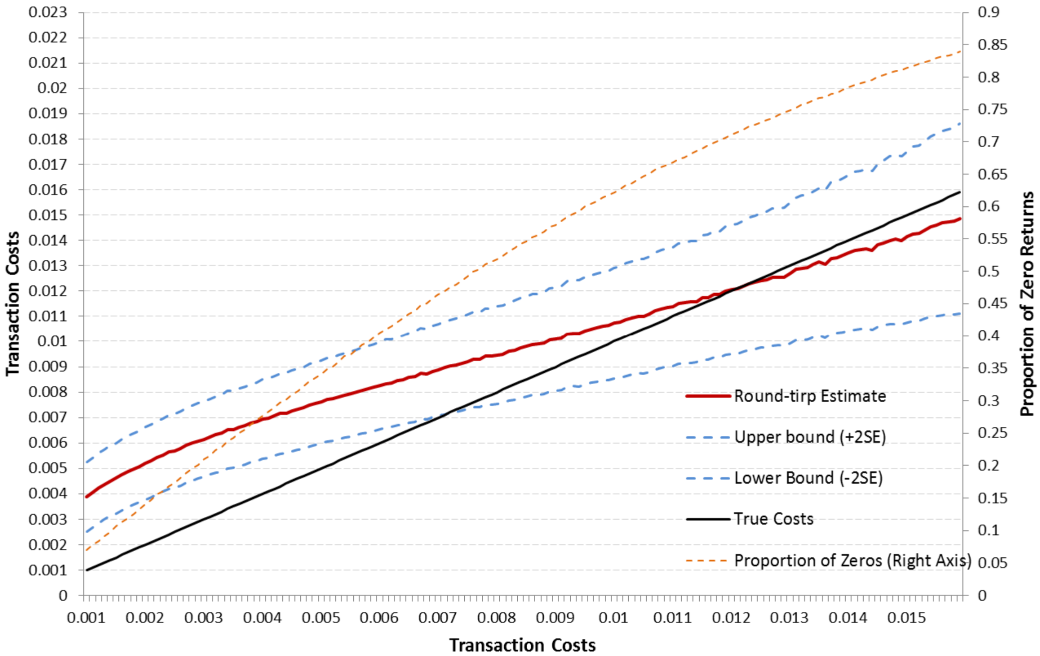
The occurrence of company specific shocks prevailing over the market-wide information depends on the ratio of idiosyncratic volatility over the true transaction cost, as it determines the probability of idiosyncratic risks to overcompensate for the opposite market movement and the one-way transaction cost.

We illustrate the nature of the bias simulating a feasibly calibrated latent return model, described in equation 3 and subsequently estimating transaction costs according to equation 5. Thereby we model market return series as a constant mean process with normally distributed disturbance term. The true latent return R^*_{jt} follows the market model with normally distributed disturbance term and suppressed intercept. The observed return R_{jt} is calculated according to Equation 3, lines 2-4. For each set of parameters we generate 1000 series with 350 observations each and estimate transaction costs for each run, and then average the estimates. Subsequently one of the parameters is changed by a small step size and the procedure is repeated for the new parameter set.

The results for different values of transaction costs keeping idiosyncratic volatility fixed are summarized in Figure 1. 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 becomes insignificant when the ratio of idiosyncratic volatility to transaction costs becomes close to one third (0.0025/0.007) and the frequency of observed zero returns exceeds 45%.

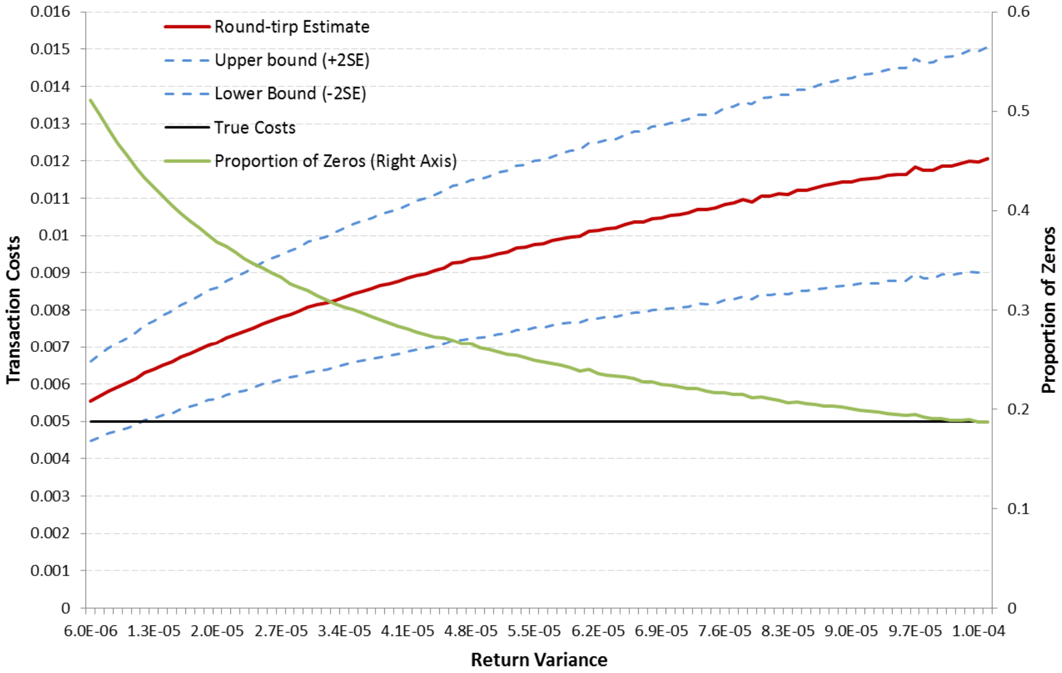
Similar picture evolves in Figure 2, where simulation results are presented for varying idiosyncratic volatility while keeping the transaction costs fixed. The overshooting bias grows monotonically and becomes significant when the proportion of zero returns falls below 40%. The estimated transaction costs become finally 2.5 times larger than the true ones, when the idiosyncratic volatility to transaction costs ratio becomes two (0.01/0.005) at the right edge of the graph.

Figure 1. Simulation results: LOT estimates vs. true transaction costs at different true transaction costs levels



The latent ('true') return follows the process $R_t^* = 0.92R_{mt} + 0.0025z_t$, $z_t \sim N(0,1)$, $R_{mt} \sim N\left(2.5\times10^{-4}, 8.3\times10^{-6}\right)$. Solid black line represents exogenously given true transaction costs, for each data point 1000 series were generated and the red line represents 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 2 . 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 an increment of 10^{-6}. This yields 100 steps, for each we estimate the model 1000 times for 350 observations.

III. Modified measure and simulation results

To address the bias described in the previous section, we develop the model that treats returns made on market information and idiosyncratic shocks separately. Stock *j* return is now defined over 5 regions:

$$\begin{aligned}
&R_{jt}^{*} = \beta_j R_{mt} + \epsilon_{jt} \\
&R_{jt} = R_{jt}^{*} - \alpha_{1j}, \quad for\ \Lambda_1 = \{If\ R_{mt} < 0\ and\ R_{jt} < 0\} \\
&R_{jt} = R_{jt}^{*} - \alpha_{2j}, \quad for\ \Lambda_2 = \{If\ R_{mt} > 0\ and\ R_{jt} > 0\} \\
&R_{jt} = R_{jt}^{*} - \alpha_{1j}, \quad for\ \Lambda_3 = \{If\ R_{mt} > 0\ and\ R_{jt} < 0\} \\
&R_{jt} = R_{jt}^{*} - \alpha_{2j}, \quad for\ \Lambda_4 = \{If\ R_{mt} < 0\ and\ R_{jt} > 0\} \\
&R_{jt} = 0, \qquad\qquad for\ \Lambda_0 = \{If\ \alpha_{1j} < R_{jt}^{*} < \alpha_{2j}\}
\end{aligned}, \tag{4}$$

whereby with regions Λ_3 and Λ_4 we treat the earlier 'problem areas'. 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 a distribution. Maintaining the assumption of idiosyncratic shocks we employ the Fisher-Tippet-Gnedenko theorem to make an inference about the distribution:

Theorem: Let $\left(X_1, X_2, ..., X_n\right)$ be a sequence of i.i.d. random variables, and $M_n = \max\left\{X_1, X_2, ..., X_n\right\}$. If $\exists\left(a_n, b_n\right): a_n > 0$, and $\lim_{n\to\infty} P\left(\left(M_n - b_n\right)/a_n \le x\right) = F\left(x\right)$, and $F\left(x\right)$ 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:

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

where $\varepsilon_{3,4}$ are the residuals from the corresponding equations of system (4). 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:

$$\begin{aligned} L(\alpha_{1j}, \alpha_{2j}, \beta_j, \sigma_j \mid R_{jt}, R_{mt}) = \prod_{\Lambda_0} \left(\Phi_{2j} - \Phi_{1j}\right) \times \\ \times \prod_{\Lambda_1} \frac{1}{\sigma_j} \phi\left(R_{jt} + \alpha_{1j} - \beta_j R_{mt}\right)^2 \prod_{\Lambda_2} \frac{1}{\sigma_j} \phi\left(R_{jt} + \alpha_{2j} - \beta_j R_{mt}\right)^2 \times \quad (5) \\ \times \prod_{\Lambda_3} f_{min}\left(R_{jt} + \alpha_{1j} - \beta_j R_{mt}\right) \prod_{\Lambda_4} f_{max}\left(R_{jt} + \alpha_{2j} - \beta_j R_{mt}\right) \end{aligned}$$

The Log-likelihood is:

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

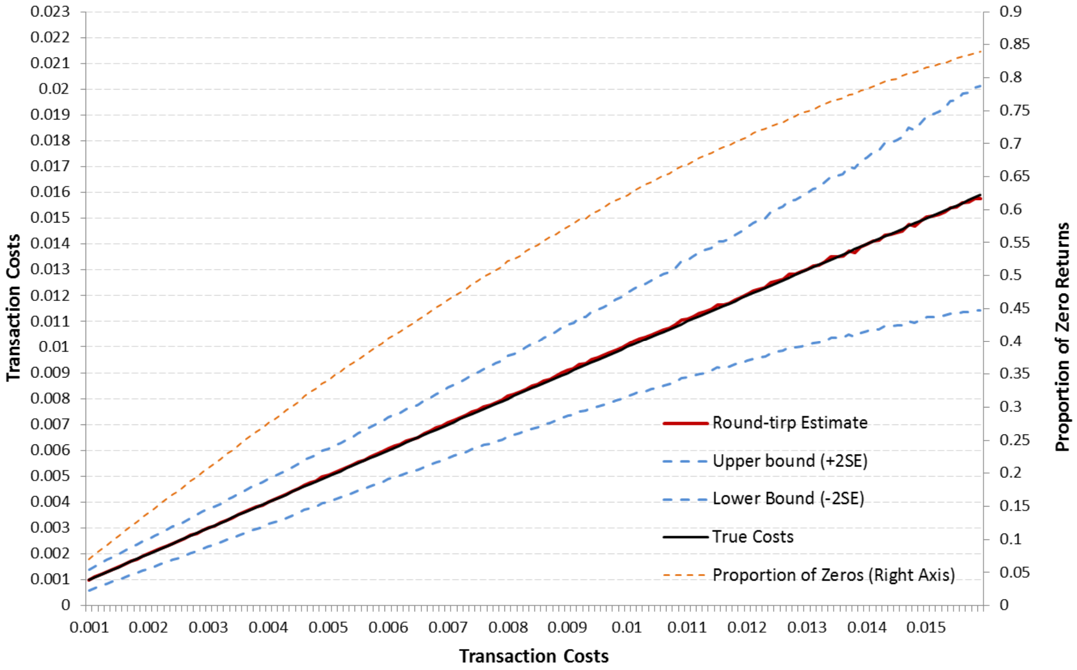
(6)

The Gumbel distribution CDF is logarithmically concave, hence optimization of the equation (6) yields the global maximum.

Turning to simulations, we see that for the modified measure in the same parameter range as for LOT the bias disappears and we obtain a very precise estimate, independent of true parameter value (Figure 3) or idiosyncratic variance (Figure 4).

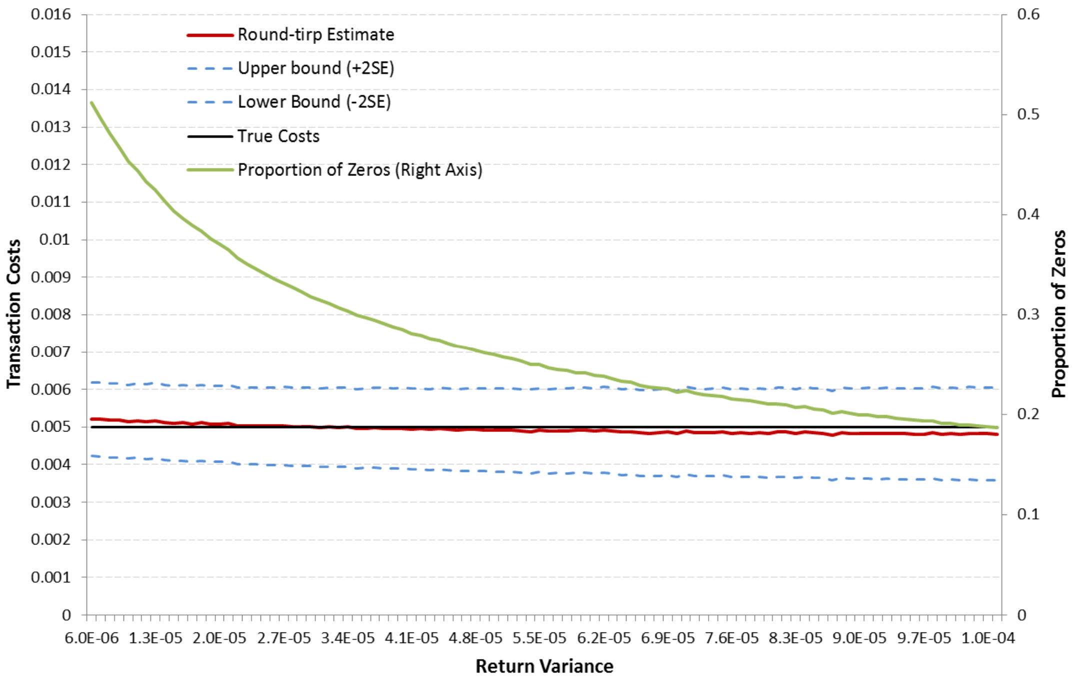
We also examine robustness of our results to alternative residual distributions. In order to capture the effect of leptokurtic distributions often observed on the real data we utilize Student's *t*-distribution with 5, 8, 10, and 12 degrees of freedom. As one can see from the Figures 5-8 the GEV estimate performs better than the LOT estimate 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. That is, even though with decreasing liquidity the bias of our measure increases, for all sensible liquidity levels it is still preferable to the LOT measure.

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



The latent ('true') return follows the process $R_t^* = 0.92R_{mt} + 0.0025z_t$, $z_t \sim N(0,1)$, $R_{mt} \sim N\left(2.5\times10^{-4}, 8.3\times10^{-6}\right)$. Solid black line represents exogenously given true transaction costs, for each data point 1000 series were generated and the red line represents 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 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 an increment of 10^{-6}. This yields 100 steps, for each we estimate the model 1000 times for 350 observations.

Figure 5. GEV vs. LOT estimates at different transaction costs levels, t-dist. with 12 d.f.

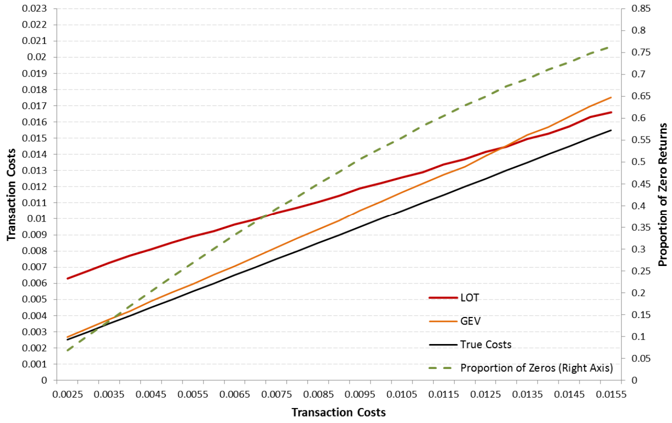


Figure 6. GEV vs. LOT estimates at different transaction costs levels, t-dist. with 8 d.f..

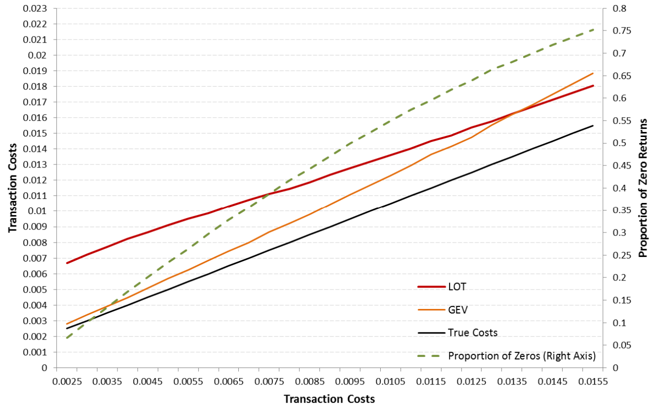


Figure 7. GEV vs. LOT estimates at different transaction costs levels, t-dist. with 10 d.f.

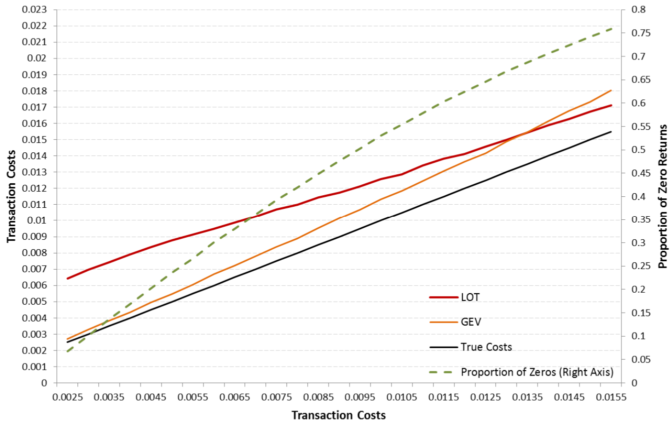
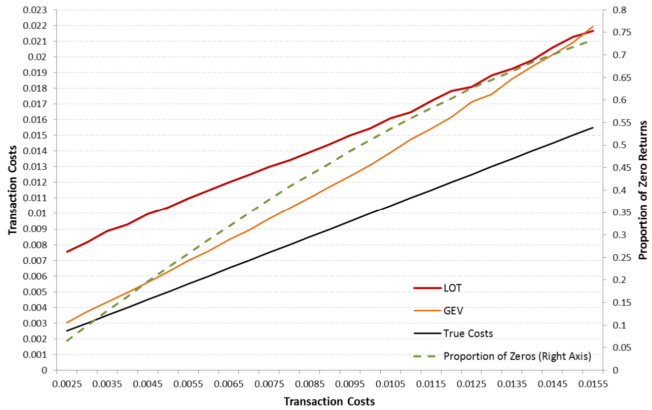


Figure 8. GEV vs. LOT estimates at different transaction costs levels, t-dist. with 5 d.f..



The return process is the same as in Figures 1 and 3, 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 is 0.0005, so that there are 30 steps for each distribution choice.

III. Horserace of LOT and GEV measures on S&P 500 constituents' stock price data

We also evaluate the proposed estimation approach performing an empirical check for a sample of the US stock market data, since the US data is considered to be standard object of empirical analysis. The sample includes closing transaction, bid and ask prices for stocks included in S&P 500 starting from January 1st 1993 to December 31st 2011. The data was obtained from the Center for Research in Securities Prices.

Following Lesmond (2005) and Amihud (2002) we filter data dropping stock-years with less than 120 observations; daily returns greater than 50% by absolute value; and proportional spread values that are greater than 80% or less than zero (considering this values as the CRSP entry mistakes).

Table 1 Panel A summarizes statistics for the two transaction costs measures as well as of our benchmark –quoted spread - for the full sample, as well as for two subsamples, divided by the major tick size reform: the decimal price increment era (2002 – 2011), and to the 1/8th , 1/16th era (1993 – 2000).[1] Average and median quoted spreads for the whole sample are 59 and 21 basis points correspondingly, this is about ten times lower, than reported by Goyenko et al (2009, Table 1) for the random sample of the US stocks, and corresponds to values reported for Dow Jones stocks in 1990-s (Goyenko et al 2009, Fig. 1). The LOT measure significantly exceeds the proportional spread estimates both in mean and median. The proposed GEV measure estimates are lower in mean than the quoted spread.

Table 1 Panel B reports the average cross-sectional correlation of benchmark and LDV estimates, as well as time series correlation of the equally weighted portfolio. The average cross-sectional GEV correlations with the benchmark are higher than those of LOT, these difference is significant on the 5% level for the whole sample and on the 10% level for both pre-decimal and decimal era sub-periods. Time series correlations also reported in Panel B are high for both measures, but statistically indistinguishable from one another. The cross-sectional correlations by year are represented in Table 2. Overall, the correlation of indirect measures with benchmark decline over time. For years 1999, 2000 and 2009 GEV significantly outperforms LOT, LOT does not significantly outperform GEV in any year.

---

[1] Both LOT and GEV models may not achieve convergence. Such cases correspond to the small number of zero return days. For example Table A6 in Appendix represents the descriptive statistics for bid-ask spread and proportion of zero return days subsamples across which the indirect measures did not converge while the spread quotes were available. The maximum proportion of zero return days corresponds to 7-8 days without price movement per year; and the mean corresponds to 2 zero return days for GEV.

Following Goyenko et al. (2009) we construct the series of deviations of each measure from spread (see Table 1 Panel C). We reveal as expected positive bias for LOT. For our modified measure we report negative average bias, it is much smaller in absolute value than that of LOT for the decimal era and larger for the 1993-2000 subperiod. We also report root mean squared prediction error which is lower for GEV in the full sample. The comparative performance in subsamples varies: in 1993-2000, where tick sizes were measured in 1/8ths and 1/16ths of a dollar our measure has a slightly larger squared error, however in the decimal era subsample (2002-2011) GEV measure is about three times more precise than the LOT measure.

To sum up, we show that our modification better captures cross-sectional variation of liquidity among S&P 500 stocks than LOT, as it has higher correlation with the benchmark. Moreover, our measure is more precise on this sample, as it has a smaller squared prediction error.

Table (1). Descriptive statistics, Correlations, Accuracy

	1993 - 2011			2002 – 2011 Decimal Era			1993 – 2000 1/8th, and 1/16th Era		
Panel A	*Spread*	*LOT*	*GEV*	*Spread*	*LOT*	*GEV*	*Spread*	*LOT*	*GEV*
Mean	0.005942	0.009384	0.002162	0.001553	0.004712	0.000583	0.012356	0.016685	0.004753
Median	0.002072	0.006585	0.000798	0.000836	0.003863	0.000396	0.010390	0.014007	0.003430
Maximum	0.109456	0.106389	0.071317	0.038710	0.071000	0.014359	0.109456	0.106389	0.071317
Minimum	0.000119	1.69E-09	4.17E-08	0.000119	1.69E-09	4.17E-08	0.000440	3.14E-09	5.57E-08
Std. dev.	0.007901	0.009427	0.003716	0.002458	0.004480	0.000800	0.009023	0.010627	0.003943
Obs.	8028	7953	7948	4603	4561	4563	3001	2972	2967
Panel B									
Cross-sectional correlation		0.570056	0.600164†		0.453941	0.483764*		0.729212	0.760826*
(SE of average correlation)		(0.00884)	(0.00860)		(0.01302)	(0.01272)		(0.01238)	(0.01180)
Time Series correlation		0.977102	0.875547		0.932227	0.768332		0.825434	0.519098
(SE)		(0.05160)	(0.11718)		(0.16900)	(0.22629)		(0.23046)	(0.34894)
Panel C									
Mean Bias		0.003481	-0.003735		0.003158	-0.000961		0.004433	-0.007527
Median Bias		0.002837	-0.001177		0.002685	-0.000398		0.003740	-0.006242
Root MSPE		0.006556	0.005270		0.004936	0.001554		0.008538	0.009444

† indicates significant difference on the 5% level and * - on the 10% level

Table (2). Cross-sectional correlations and standard errors in percent

Year	*1993*	*1994*	*1995*	*1996*	*1997*	*1998*	*1999*	*2000*	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*
LOT	88.9	87.6	77.2	73.4	61.7	64.3	68.4	61.9	45.8	59.8	48.9	50.2	49.4	39.6	20.5	18.4	65.9	49.2	52.2
SE	2.5	2.6	3.4	3.6	4.1	3.9	3.7	3.9	4.4	3.9	4.2	4.1	4.1	4.3	4.5	4.5	3.5	4.0	4.0
GEV	87.5	83.7	76.7	76.4	67.6	66.6	77.6*	72.6†	47.9	63.4	53.0	45.5	55.8	43.0	19.8	26.2	74.9*	57.7	44.6
SE	2.7	3.0	3.5	3.4	3.8	3.8	3.2	3.4	4.3	3.8	4.1	4.3	3.9	4.2	4.6	4.5	3.1	3.7	4.2

† indicates significant difference on the 5% level and * - on the 10% level

Conclusion

We explain an upward bias in a wide-spread measure of transaction costs – Lesmond et al. (1999, LOT) – through the mistreatment of large idiosyncratic shocks, counter directional to the market. We introduce a modification which significantly outperforms LOT in simulations and on standard US data. This makes the use of our measure extremely advantageous for estimating liquidity on rather liquid markets with high idiosyncratic volatility.

References

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

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

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