Dealing with Negative Values in the High-Low Spread Estimator

A Comment On
"Pricing of Liquidity Risks: Evidence from Multiple Liquidity Measures"
by Soon-Ho Kim and Kuan-Hui Lee, *Journal of Empirical Finance* (25), 2014, 112-133.

Shane A. Corwin
March 2014

In their recent paper in the *Journal of Empirical Finance*, Kim and Lee (2014) examine the pricing of liquidity risk using several alternative measures of liquidity. Their analysis focuses on eight liquidity measures, including the high-low spread measure developed in Corwin and Schultz (CS 2012). In this note, I use the summary statistics from Kim and Lee to highlight an adjustment that should be made when applying the CS measure. The adjustment has to do with negative daily values of the estimator.[1]

As noted in Corwin and Schultz (2012), the estimated high-low spread can be negative for some two-day periods. In these cases, we suggest making an adjustment for negative daily values before taking the monthly average of the CS measure.[2] Several adjustments are possible, including setting the negative values to zero, treating them as missing, or possibly setting them equal to the minimum tick size. In our comparison to TAQ data, we evaluate the first two methods and find that setting negative daily values to zero produces monthly averages that most closely match TAQ spreads. This adjustment is utilized both in the CS estimates we provide on our webpage and in the sample program we provide for estimating the CS measure.

The adjustment to the CS measure described above is similar in spirit to the common correction for positive serial correlations when using the Roll estimator (see Roll (1984) and Harris (1990)).[3] Roll provides an estimate for the percentage spread defined as $2 \cdot \sqrt{-COV}$, where *COV* is the serial covariance in daily returns. The Roll measure cannot be computed when the serial covariance is positive. In these cases, researches typically make one of three modifications: (1) treat the Roll measure as missing, (2) set the Roll measure to zero, or (3) take the square root without applying the negative sign and treat the result as a negative spread. The second approach is most common and is applied in both Corwin and Schultz (2012) and Goyenko, Holden, and Trzcinka (2009). The third approach is applied in the original Roll (1984) paper. In contrast, both Kim and Lee (2014) and Lesmond (2005) treat positive covariances as if they are negative, resulting in a positive Roll spread estimate.

To demonstrate the potential impact of these adjustments, I examine and reestimate the summary statistics provided by Kim and Lee (2014). While these summary statistics are not appropriate as a test of

[1] This note is not intended as a criticism of the results in Kim and Lee. In fact, it is possible that the adjustment suggested here would have very little impact on their primary conclusions. I use their paper only to illustrate the suggested adjustments for negative daily values of the CS measure.

[2] While the CS measure can be averaged across different time periods, Corwin and Schultz (2012) test the performance of the measure using monthly estimates. See www.nd.edu/~scorwin for applications based on daily and intraday analysis. The web page also provides links to a sample program for estimating the CS measure and to monthly CS estimates for all CRSP securities from 1926 through the most recent year.

[3] By construction, the LOT measure cannot be negative due to the bounds imposed in the maximum likelihood estimator. However, it is not uncommon to obtain zero estimates for the buying cost, the selling cost, and the overall LOT measure.

alternative spread estimators, they do serve to illustrate the importance of applying the adjustments described above.[4] In their analysis, Kim and Lee do not apply an adjustment for negative daily values of the CS measure. The impact of the negative values is evident in their Table 1, where they find negative average values for the CS measure, especially for large stocks. The impact can also be seen in Table 2, where they report a negative time series correlation between a market-wide average of the CS measure and market-wide averages of several other liquidity measures. Below I present similar summary statistics obtained when I reestimate their results both without an adjustment for negative daily CS estimates and with one of two adjustments: (1) setting negative values to missing, or (2) setting negative values to zero. For completeness, I also provide results for three different versions of the Roll spread that differ in their treatment of cases with positive serial covariance; The first treats these cases as zero spreads (as in Corwin and Schultz (2012) and Goyenko et al. (2009), the second treats the cases as negative spreads (as in Roll (1984)), and the third treats these cases as positive spreads (as in Lesmond (2005) and Kim and Lee (2014)). Finally, I provide results based on the two alternative versions of the LOT estimator described in Goyenko et al. (2009): LOT Mixed and LOT Y-Split.

Although I do not perfectly replicate their results, my sample formation process is chosen to closely match that in Kim and Lee (2014). I begin with all NYSE- and Amex-listed common stocks in CRSP for the period from July 1962 through December 2012.[5] I exclude stocks that trade on fewer than 100 days during this sample period. In addition, to be included in a particular month, a stock must have a positive market value and at least 12 days of high-low price data. I eliminate all observations for a stock during a given year if the stock experiences a stock split or is delisted during the year, or if the stock price was below $5 or above $1,000 at the end of the prior year. While Kim and Lee examine eight liquidity measures, I provide results only for the CS measure, the Roll (1984) measure, the Amihud (2002) measure, and the LOT measure (Lesmond et al. (1999)) measure, including multiple versions of some measures. I eliminate stock–months during which any of the estimators is missing. Finally, during each year, I assign stocks to one of 25 market value categories based on market capitalization at the end of the prior year.[6]

Table I lists the equally weighted average for each liquidity measure across all stock-months for each of the 25 size-based portfolios. These summary statistics are similar to those presented in Table 1 of Kim and Lee (2014). I focus first on the alternative versions of the Corwin-Schultz High-Low Spread measure. When Ie apply no adjustment for negative daily values in the CS measure (CS_{NoAdj}), the results are consistent with Kim and Lee (2014). While the average CS measure tends to be larger for small stocks than large stocks, the average is often close to zero and is negative for many of the large stock portfolios. When Ie apply an adjustment for negative daily values, however, the results are quite different. Regardless of whether the negative daily CS values are excluded or set to zero, the mean values of the CS measure are positive for all portfolios and decrease monotonically with firm size. Turning to the Roll measure, we see that changing the sign on positive serial covariances ($Roll_{SignChg}$) leads to an upward bias in the measure relative to alternative corrections that are more commonly applied in the literature. Mean Roll spreads are roughly twice as large when positive serial covariances are treated as negative than when these cases are set to zero. Finally, as in Goyenko et al. (2009), I find that the two versions of the LOT

[4] Corwin and Schultz (2012) and Goyenko et al. (2009) examine the performance of various low frequency spread estimators.
[5] Specifically, I include only those securities with a CRSP share code of 10, 11, or 12, and a CRSP exchange code equal to 1 or 2.
[6] These sample restrictions differ slightly from those applied in Kim and Lee (2014). Most importantly, Kim and Lee exclude financials. They also require 15 daily observations within the month and require data for all eight of their liquidity measures. I do not expect these differences in sample selection to have an important effect on the conclusions herein.

measure produce quite different results. Means for LOT Mixed range from 0.010 to 0.042 across the size categories, compared to a range of 0.003 to 0.022 for LOT Y-Split.

In Table II, I provide correlations among the various liquidity measures. Following Kim and Lee (2014), I begin by calculating time-series correlations based on market-wide averages of each liquidity measure, where market-wide averages are computed each month based on all stocks with available data. These results are presented in Panel A. For completeness, Panel B lists pooled correlations based on all cross-sectional and time-series observations. Again, the results based on the uncorrected CS measure (CS_{NoAdj}) and $Roll_{SignChg}$ are similar to those presented in Kim and Lee (2014), with negative time series correlations between a market-wide average of the CS measure and market-wide averages of some other liquidity measures. In contrast, the adjusted values of the CS measure tend to have large positive correlations with other liquidity measures. These findings highlight the importance of applying appropriate adjustments for negative values when computing the CS measure. The results also illustrate the potential impact of alternative corrections to the Roll measure, as correlations with other liquidity measures differ markedly depending on the treatment of positive serial covariances.

Summary

This note highlights the importance of adjustments for negative daily values when utilizing the High-Low Spread estimator developed by Corwin and Schultz (2012). It also compares alternative versions of the Roll (1984) spread and the LOT (1999) measure. See Corwin and Schultz (2012) for a complete description of these and other potential adjustments for empirical application of the High-Low Spread measure, including adjustments for overnight returns and non-trading.

References

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31-56.

Corwin, Shane A., and Paul Schultz, 2012, A simple way to estimate bid-ask spreads from daily high and low prices, *Journal of Finance* 67, 719-759.

Goyenko, Ruslan Y., Craig W. Holden, and Charles A. Trzcinka, 2009, Do liquidity measures measure liquidity, *Journal of Financial Economics* 92, 153-181.

Harris, Lawrence, 1990, Statistical properties of the Roll serial covariance bid/ask spread estimator, *Journal of Finance* 45, 579-590.

Lesmond, David, Joseph Ogden, and Charles Trzcinka, 1999, A new estimate of transactions costs, *Review of Financial Studies* 12, 1113-1141.

Kim, Soon-Ho, and Kuan-Hui Lee, Pricing of liquidity risks: Evidence from multiple liquidity measures, *Journal of Empirical Finance* 25, 112-133.

Roll, Richard, 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance* 39, 127-1139.

Table I – Summary Statistics for Illiquidity Measures by Size Portfolio

The table reports mean values for market value, monthly return, and several liquidity measures across all stock-months that meet the sample restrictions described in the preceding text. Stocks are sorted into 25 size categories at the end of each calendar year based on market capitalization at the end of the year.

Portfolio	Average N	Market Cap	Monthly Return	Amihud	LOT Mixed	LOT Y-Split	Corwin-Schultz High-Low Spread			Roll Spread		
							Adjustment for Negative Values			Adjustment for Positive Autocov.		
							No Adj.	Zero	Missing	Zero	Neg. Sprd	Sign Chg.
Small	46.10	24,295	0.0143	4.0699	0.0421	0.0222	0.0062	0.0136	0.0189	0.0132	0.0055	0.0210
2	45.84	52,593	0.0136	2.3196	0.0389	0.0196	0.0034	0.0114	0.0166	0.0133	0.0054	0.0212
3	45.86	83,967	0.0117	1.6405	0.0362	0.0180	0.0026	0.0107	0.0158	0.0133	0.0056	0.0210
4	45.30	117,938	0.0119	1.1935	0.0343	0.0164	0.0019	0.0099	0.0150	0.0129	0.0053	0.0206
5	45.28	154,385	0.0120	0.9479	0.0322	0.0153	0.0015	0.0093	0.0142	0.0122	0.0046	0.0199
6	45.70	196,660	0.0117	0.7081	0.0305	0.0144	0.0013	0.0089	0.0137	0.0119	0.0045	0.0193
7	45.35	248,870	0.0119	0.5714	0.0287	0.0130	0.0007	0.0082	0.0130	0.0112	0.0038	0.0185
8	45.20	291,830	0.0096	0.4900	0.0277	0.0124	0.0006	0.0080	0.0127	0.0109	0.0036	0.0183
9	45.27	352,032	0.0103	0.3912	0.0262	0.0115	0.0004	0.0077	0.0123	0.0105	0.0032	0.0178
10	45.04	416,436	0.0103	0.3125	0.0255	0.0110	0.0003	0.0075	0.0121	0.0105	0.0033	0.0176
11	44.84	494,740	0.0094	0.2768	0.0238	0.0100	0.0001	0.0072	0.0116	0.0100	0.0030	0.0170
12	44.98	585,100	0.0099	0.2282	0.0228	0.0096	0.0001	0.0070	0.0114	0.0095	0.0025	0.0165
13	45.33	690,622	0.0098	0.1866	0.0219	0.0089	0.0000	0.0067	0.0110	0.0091	0.0023	0.0160
14	44.83	818,270	0.0096	0.1597	0.0208	0.0083	0.0000	0.0066	0.0109	0.0090	0.0025	0.0156
15	44.98	950,076	0.0099	0.1378	0.0201	0.0080	0.0001	0.0064	0.0105	0.0087	0.0023	0.0151
16	44.76	1,126,200	0.0096	0.1173	0.0191	0.0074	-0.0001	0.0062	0.0103	0.0085	0.0022	0.0149
17	44.38	1,328,001	0.0094	0.0934	0.0182	0.0069	-0.0002	0.0061	0.0102	0.0081	0.0017	0.0146
18	44.01	1,611,125	0.0092	0.0734	0.0176	0.0067	-0.0002	0.0059	0.0099	0.0078	0.0014	0.0143
19	44.20	2,012,847	0.0090	0.0559	0.0168	0.0061	-0.0004	0.0059	0.0099	0.0077	0.0011	0.0142
20	44.06	2,530,386	0.0096	0.0396	0.0157	0.0057	-0.0003	0.0057	0.0096	0.0075	0.0013	0.0137
21	44.57	3,418,631	0.0089	0.0300	0.0148	0.0052	-0.0003	0.0056	0.0094	0.0071	0.0008	0.0133
22	44.59	4,592,934	0.0083	0.0227	0.0140	0.0048	-0.0002	0.0055	0.0092	0.0068	0.0008	0.0129
23	43.90	6,251,080	0.0077	0.0180	0.0134	0.0045	-0.0002	0.0053	0.0091	0.0065	0.0005	0.0126
24	43.28	9,957,149	0.0074	0.0126	0.0120	0.0039	-0.0002	0.0053	0.0091	0.0065	0.0006	0.0123
Large	44.00	34,664,018	0.0073	0.0056	0.0100	0.0030	-0.0002	0.0050	0.0086	0.0060	0.0003	0.0116

Notes:
The liquidity measures are the Amihud (2002) measure, the LOT measure (Lesmond, Ogden, and Trzcinka (1999), the Roll (1984) measure, and the Corwin and Schultz (CS 2012) High-Low Spread estimator. The two variations of the LOT measure (LOT Mixed and LOT Y-Split) differ in how they divide the three regions of the maximum likelihood function (see Goyenko et al. (2009)). The three versions of the CS measure differ in how they handle negative daily values of the estimator when computing monthly averages (see Corwin and Schultz (2012)). The three alternatives are: (1) no adjustment, (2) setting negative daily values to zero, and (3) treating negative daily values as missing. The three versions of the Roll measure differ in how they adjust for positive serial covariances in the formula $2 \cdot \sqrt{-COV}$. The three alternatives are: (1) defining the Roll measure as zero, (2) calculating the square root without the negative sign and defining the resulting Roll measure as negative, and (3) changing the sign of the covariance and defining the Roll measure as positive.

Table II – Correlations among Illiquidity Measures

The table reports correlations among several monthly liquidity measures. The sample includes all stock-months that meet the sample restrictions described in the preceding text. Panel A reports time-series correlations among market-wide averages of the liquidity measures and Panel B reports pooled correlations based on all stock-months with available data. The liquidity measures are the Amihud (2002) measure, the LOT measure (Lesmond, Ogden, and Trzcinka (1999), the Roll (1984) measure, and the Corwin and Schultz (CS 2012) High-Low Spread estimator. The two variations of the LOT measure (LOT Mixed and LOT Y-Split) differ in how they divide the three regions of the maximum likelihood function (see Goyenko et al. (2009)). The three versions of the CS measure differ in how they handle negative daily values of the estimator when computing monthly averages (see Corwin and Schultz (2012)). The three alternatives are: (1) no adjustment, (2) setting negative daily values to zero, and (3) treating negative daily values as missing. The three versions of the Roll measure differ in how they adjust for positive serial covariances in the formula $2 \cdot \sqrt{-COV}$. The three alternatives are: (1) defining the Roll measure as zero, (2) calculating the square root without the negative sign and defining the resulting Roll measure as negative, and (3) changing the sign of the covariance and defining the Roll measure as positive.

	Amihud	LOT Mixed	LOT Y-Split	Corwin-Schultz High-Low Spread			Roll Spread	
				Adjustment for Negative Values			Adj. for Pos. Autocov.	
				No Adj.	Zero	Missing	Zero	Neg. Sprd
Panel A – Time Series Correlations among Market-Wide Illiquidity Measures								
LOT Mixed	0.681							
LOT Y-Split	0.660	0.947						
CS_{NoAdj}	0.067	-0.121	-0.048					
CS_{Zero}	0.632	0.236	0.122	0.194				
CS_{Miss}	0.294	0.174	0.042	0.028	0.981			
$Roll_{Zero}$	0.116	-0.018	-0.121	0.267	0.823	0.811		
$Roll_{NegSprd}$	-0.073	-0.244	-0.265	0.590	0.436	0.374	0.811	
$Roll_{SignChg}$	0.252	0.198	0.053	-0.124	0.906	0.945	0.839	0.362
Panel B – Pooled Correlations among Illiquidity Measures								
LOT Mixed	0.326							
LOT Y-Split	0.275	0.684						
CS_{NoAdj}	0.311	0.166	0.222					
CS_{Zero}	0.446	0.457	0.377	0.655				
CS_{Miss}	0.393	0.439	0.325	0.399	0.933			
$Roll_{Zero}$	0.160	0.251	0.155	0.355	0.470	0.439		
$Roll_{NegSprd}$	0.096	0.079	0.063	0.457	0.294	0.204	0.875	
$Roll_{SignChg}$	0.172	0.383	0.213	-0.015	0.482	0.566	0.624	0.167