Price Impact or Trading Volume: Why Is the Amihud (2002) Measure Priced?

Xiaoxia Lou
Alfred Lerner College of Business, University of Delaware

Tao Shu
Terry College of Business, University of Georgia
HKUST Business School

The return premium associated with the Amihud (2002) measure is generally considered a liquidity premium that compensates for price impact. We find that the pricing of the Amihud measure is not attributable to the construction of the return-to-volume ratio intended to capture price impact, but is driven by the trading volume component. Additionally, the high-frequency price impact and spread benchmarks are priced only in January and do not explain the pricing of the trading volume component of the Amihud measure. Additional analyses suggest that the volume effect on stock return is likely caused by mispricing, not by compensation for illiquidity. (*JEL* G10, G12)

Received September 20, 2014; editorial decision April 16, 2017 by Editor Andrew Karolyi.

The Amihud (2002) measure is one of the most widely used liquidity proxies in the finance literature.[1] From 2009 to 2015, more than 120 papers published in the *Journal of Finance*, the *Journal of Financial Economics*, and the *Review of Financial Studies* have used the Amihud measure in their empirical analyses.[2] The Amihud measure has three advantages over many other liquidity measures. First, the Amihud measure has a simple construction that uses the absolute value of the *daily* return-to-volume ratio to capture price

The paper greatly benefited from the comments of two anonymous referees and Andrew Karolyi (the editor). We appreciate the helpful comments from Yakov Amihud, Michael Brennan, Kalok Chan, Darwin Choi (CICF discussant), Sudipto Dasgupta, Amit Goyal, Joel Hasbrouck (AFA discussant), Jennifer Huang, Ronnie Sadka, Johan Sulaeman, Sheridan Titman, Liyan Yang, Tong Yao, Yuan Yu, Bohui Zhang, and Chu Zhang and the participants at the 2015 American Finance Association Meeting, 2014 China International Conference of Finance, University of South Carolina, HKUST, University of Hong Kong, Chinese University of Hong Kong, Shanghai Advanced Institute of Finance (SAIF), City University of Hong Kong, Fudan University, and Zhejiang University. A substantial amount of work for this paper was conducted when Shu was visiting HKUST Business School. Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Xiaoxia Lou, University of Delaware, Newark, DE 19716; telephone: 302-831-8773. E-mail: lous@udel.edu.

[1] Besides the Amihud (2002) measure as a price impact (cost-per-dollar-volume) proxy, the finance literature also has proposed many measures for the three aspects of liquidity: spread, price impact, and resilience (see Holden, Jacobsen, and Subrahmanyam 2014 for a survey).

[2] Note that we only count published papers and exclude any forthcoming papers.

doi:10.1093/rfs/hhx072 Advance Access publication July 11, 2017

impact. Second, the measure uses daily data and therefore provides a longer time series relative to intra-daily proxies based on TAQ data. Third, the measure has a strong positive relation with expected stock return (Amihud 2002; Chordia, Huh, and Subrahmanyam 2009, among many other studies). The positive return premium of the Amihud measure is generally considered a liquidity premium that compensates for price impact.

Theoretically, however, it is unclear that the Amihud measure would be priced because of the compensation for price impact. As discussed by Chordia, Huh, and Subrahmanyam (2009), "Although many microstructure theories have been developed, extant economic models are unable to map precisely onto the Amihud (2002) construct of the ratio of absolute return to volume" (p. 3630). Because the Amihud measure is widely used to examine liquidity premium or control for liquidity, it is important to know whether the pricing of the Amihud measure is indeed caused by price impact or by other reasons. Furthermore, examining the pricing of the Amihud measure also helps us understand liquidity measurement and liquidity premium. For example, the return premium of the Amihud measure is generally considered to be direct evidence that investors, as predicted by theory, demand compensation for price impact or transaction cost.

This paper studies the pricing of the Amihud (2002) measure from a new perspective, the close connection between the Amihud measure and trading volume, as illustrated by the construction of the measure:

$$A_{it} = \frac{1}{D_{it}} \sum_{d=1}^{Dit} \frac{|r_{id}|}{Dvol_{id}}, \tag{1}$$

where A_{it} is the Amihud measure of firm i estimated in month t; r_{id} and $Dvol_{id}$ are daily return and daily dollar trading volume for stock i on day d; and D_{it} is the number of days with available ratio in month t.[3] With everything else equal, higher trading volume leads to a lower Amihud measure.[4] This linkage is particularly strong because the trading volume component has a much greater cross-sectional variation than does the stock return component. For example, the 75th percentile cutoff of the trading volume component is over 100 times its 25th percentile cutoff, but the 75th percentile cutoff of the return component is just twice its 25th percentile cutoff.[5]

[3] Amihud (2002) constructs the measure annually, and existing studies use both monthly and annual measures. We use the monthly measure for the main analysis because it reflects more recent information, and we conduct robustness tests using the annual measure.

[4] Some studies further adjust the Amihud measure for inflation or trend in trading volume. The approaches of our analyses are such that we need not to do so. For sorting analysis, we sort stocks into portfolios every month. For the Fama-MacBeth regression analysis that uses the Amihud measures as independent variables, we follow the literature (e.g., Brennan, Huh, and Subrahmanyam 2013) and transform the measures into natural logs. Doing so makes the scaling irrelevant.

[5] The corresponding statistics are presented in Table 1 and discussed in Section 1.2.

To focus on the trading volume component of the Amihud measure, we construct a "constant" version of the Amihud measure, *A_C*, by replacing absolute return in the Amihud measure with one:

$$A\_C_{it} = \frac{1}{D_{it}} \sum_{d=1}^{Dit} \frac{1}{Dvol_{id}}, \tag{2}$$

where all the components are as defined in Equation (1). The *A_C* measure has a correlation of 0.90 with the original Amihud measure, suggesting that the variation in the Amihud measure is driven, in large part, by the variation in the trading volume component. Additionally, we find that the "constant" measure is priced similarly to the original measure: stocks in the top quintile of *A_C* outperform those in the bottom quintile by 0.61% (t-stat 2.95) per month in raw return and 0.44% (t-stat 3.20) in four-factor alpha that controls for the three Fama-French factors and the momentum factor. This is very close to the spread based on the original Amihud measure: 0.56% (t-stat 2.36) per month in raw return and 0.35% (t-stat 2.31) in four-factor alpha.

We further find that a residual Amihud measure, the residual from cross-sectional regressions of A on *A_C* and therefore orthogonal to the constant measure *A_C*, is not associated with a positive return premium. In fact, the top quintile of the residual measure *underperforms* the bottom quintile by 0.17% (t-stat 1.05) per month in raw return and 0.16% (t-stat 0.96) in four-factor alpha. These results indicate that the pricing of the Amihud measure is driven by its trading volume component, not by its construct of return-to-volume ratio. We reach the same conclusion using the firm-level Fama and MacBeth (1973) regressions of monthly stock returns on the Amihud measures controlling for size, book-to-market ratio, momentum, and short-term return reversal. The coefficient of the "constant" measure is significantly positive but on the residual Amihud measure it is either insignificant or significantly negative.

Our results are similar when we use the turnover-based Amihud measure (*AT*) proposed by Brennan, Huh, and Subrahmanyam (2013) that is constructed using the absolute return-to-turnover ratio instead of the absolute return-to-volume ratio. The results also hold for a battery of robustness tests including using annual Amihud measures, the NASDAQ stocks, the sub-periods, the ranks instead of raw values of the independent variables, or controlling for idiosyncratic return volatility.

Because the pricing of the Amihud measure is generally considered compensation for price impact, we directly examine the role of price impact in explaining the pricing of the Amihud measure using a high-frequency price impact benchmark widely used in the literature (Hasbrouck 2009; Goyenko, Holden, and Trzcinka 2009). The price impact benchmark, λ, is constructed for NYSE/AMEX stocks from 1983 to 2012 as the slope coefficient of 5-minute stock return regressed on the signed square root 5-minute trading volume for a firm-month. We also consider an alternative non-volume-based price impact measure, the percentage 5-minute price impact (*PI*), which evaluates the permanent price change of a given trade (Goyenko, Holden, and Trzcinka 2009). We further expand the analysis to bid-ask spread and construct three widely used high-frequency spread benchmarks including percentage quoted spread (*QS*), percentage effective spread (*ES*), and percentage realized spread (*RS*) (Goyenko, Holden, and Trzcinka 2009; Fong, Holden, and Trzcinka Forthcoming).

Consistent with the existing literature (Hasbrouck 2009; Goyenko, Holden, and Trzcinka 2009), we find a correlation of 0.74 between the Amihud (2002) measure and the λ measure, which indicates that, indeed, the Amihud (2002) measure does a good job capturing price impact. However, return regression analyses show that the price impact benchmarks, either the λ measure or the *PI* measure, are not associated with a return premium or explaining the pricing of the Amihud measure. Additionally, the spread benchmarks are not associated with a return premium or explaining the pricing of the Amihud measure, either. We further decompose the Amihud measures into a transaction cost component and a non-transaction-cost component and examine their pricing separately. Specifically, we estimate cross-sectional regressions of the Amihud measures on the price impact and spread benchmarks, and calculate the transaction cost component as the fitted value of the regressions, and the noncost component as the residual of the regressions. The noncost component, therefore, is orthogonal to the price impact and spread benchmarks. The results of return regressions show that the noncost component is priced but the transaction cost component is not, indicating that the pricing of the Amihud measure is not due to its association with common liquidity benchmarks.

Because it is surprising that the liquidity benchmarks are not associated with a return premium in the full sample period, we further examine the pricing of these liquidity benchmarks. Consistent with Eleswarapu and Reinganum (1993) and Hasbrouck (2009), who show that liquidity premium is concentrated in January, we find that these liquidity benchmarks are indeed priced in January, but not in non-January months. The finding that the liquidity benchmarks are priced only in January is puzzling and unexplained by the existing theory of liquidity premium (Hasbrouck 2009).

Our results show that the pricing of the Amihud measure is due to its association with trading volume, and such pricing cannot be explained by existing liquidity benchmarks. Then what drives the pricing of trading volume? In particular, is the return premium of trading volume a liquidity premium from some dimension of liquidity that is not captured by the existing liquidity benchmarks, or is it caused by nonliquidity factors as suggested by some studies? For example, previous studies have related trading volume or its return premium to various factors such as investor disagreement (e.g., Harris and Raviv 1993; Blume, Easley, and O'Hara 1994; Kandel and Pearson 1995), value investing (Lee and Swaminathan 2000), stock visibility

(Gervais, Kaniel, and Mingelgrin 2001), information uncertainty (Jiang, Lee, and Zhang 2005; Barinov 2014), and investor sentiment (Baker and Wurgler 2006).[6]

We conduct four tests to distinguish the liquidity and nonliquidity explanations of the volume premium. Because dollar volume is the product of turnover and firm size, to remove the size effect from the volume premium, we focus on two clean measures of trading volume: (1) *AT_C*, the constant version of the turnover-based Amihud measure; and (2) *Turnover*, the monthly average of daily turnover. Both are constructed using only the turnover of a stock. The results of our four tests overall suggest that the volume premium is likely to be attributed to mispricing rather than liquidity premium.

We first examine the seasonality of the volume premium, and find that the volume premium completely disappears in January but remains strong the rest of the year. This is in stark contrast to liquidity benchmarks, which are priced in January, but not in non-January months. The timing of the pricing suggests that the underlying source of the volume premium may vastly differ from liquidity premium. Our second test is based on the notion that liquidity premium should be larger when liquidity is scarce and investors care more about stock illiquidity, such as the time periods when the aggregate liquidity is low (Pástor and Stambaugh 2003). However, contrary to this liquidity premium predication, we find that the volume premium is *not* larger after episodes of higher market illiquidity.

We also conduct two tests to explore the mispricing explanation of the volume premium. Our first test is based on Stambaugh, Yu, and Yuan (2012), who suggest that mispricing, especially overpricing, will be greater following periods of high market sentiment. We find that, consistent with the mispricing hypothesis, the volume premium is significantly larger following the high-sentiment period, and the difference is driven by the short leg, suggesting that high volume stocks are likely to be overpriced. Our second test is based on La Porta et al. (1997), who suggest that if an anomaly is associated with mispricing, then it will be stronger in the earnings announcement window, as the release of earnings helps correct mispricing.[7] We find that, consistent with this prediction, the volume premium is large and significant in the 3-day earnings announcement window but disappears in the non-announcement window. Our examination of analyst forecast errors also suggests that earnings release helps correct market overoptimism about high volume stocks relative to low volume stocks.

---

[6] Some studies also document a weak or even negative relation between volume and stock liquidity (Foster and Viswanathan 1993; Lee, Mucklow, and Ready 1993; and Johnson 2008). As another example, trading volume can be high when the markets are illiquid as was seen in the flash crash of 2010. In addition, Collin-Dufresne and Fos (2015) and Easley, Lopez de Prado, and O'Hara (2016) suggest that trading volume can capture information, such as informed trading and adverse selection, beyond the usual liquidity benchmarks.

[7] A contemporaneous study by Engelberg, McLean, and Pontiff (2016) uses this approach to study a strategy that combines 94 anomalies documented by the existing literature.

Existing literature has shown that although both characteristic liquidity and systematic liquidity are related to asset pricing, they can have different properties and behaviors (e.g., Lou and Sadka 2011; Ben-Rephael, Kadan, and Whol 2015). Although the focus of our paper is to examine the pricing of the Amihud measure with respect to characteristic illiquidity, the Amihud measure also has been used to examine the liquidity commonality (Kamara, Lou and Sadka 2008; Karolyi, Lee, and van Dijk 2012) and the pricing of liquidity as a systematic risk factor (e.g., Acharya and Pedersen 2005). Consistent with the findings in our paper, Karolyi, Lee, and van Dijk (2012) find that the commonality in turnover is the most reliable liquidity-demand-side variable to explain the time-variation in liquidity commonality where liquidity is measured by the Amihud measure. We therefore extend our analysis to the use of the Amihud measure to examine the pricing of liquidity risk (e.g., Acharya and Pedersen 2005). We construct systematic liquidity factors using the Amihud measure and its trading volume component, and conclude that the trading volume component is also primarily responsible for the pricing of the Amihud measure as a systematic factor.

1. Measure Construction and Sample Selection

1.1 Measure construction

The measures used in this paper are constructed as below:

- A: the Amihud (2002) measure, defined by Equation (1).
- A_C: the "constant" Amihud measure corresponding to A, defined by Equation (2).
- AT: the turnover-based Amihud illiquidity measure from Brennan, Huh, and Subrahmanyam (2013)

$$AT_{it} = \frac{1}{D_{id}} \sum_{d=1}^{Dit} \frac{|r_{id}|}{TO_{id}}, \tag{3}$$

where AT_{it} is the turnover-based Amihud measure for stock i in estimation month t, and TO_{id} is the turnover of stock i on day d, calculated as daily share volume divided by total shares outstanding. The other variables are as defined in Equation (1).

- AT_C: the "constant" turnover-based Amihud measure corresponding to AT

$$AT\_C_{it} = \frac{1}{D_{it}} \sum_{d=1}^{Dit} \frac{1}{TO_{id}}, \tag{4}$$

which differs from Equation (3) only in replacing the numerator of the ratio $|r_{id}|$ with a constant 1.

- $|Ret|$: return component of the Amihud measure, calculated as the monthly average of daily absolute returns over the estimation month.

We follow the literature and winsorize these measures at the 1 and 99 percentage points in each cross-section to minimize the influence of outliers. Table A1 (see the appendix) summarizes the definitions of all the variables used in the paper. In addition to the turnover-based Amihud measure, we also examine the square root version of the Amihud measure that is constructed as the Amihud (2002) measure but taking the square root of the daily absolute return-to-volume ratio. Hasbrouck (2009) proposes the square-root measure to control for skewness. We construct the "constant" measure corresponding to the square-root Amihud measure by replacing the numerator with a constant one, and repeat the tests in this paper. The results are not reported for the sake of brevity, but all our findings in this paper hold for the square-root version of the Amihud measure as well.

1.2 Sample construction

Our sample stocks include ordinary common shares (share codes 10 and 11) listed on the NYSE and the AMEX.[8] We exclude NASDAQ stocks because their trading volume is inflated relative to that of NYSE/AMEX stocks because of different trading mechanisms.[9] We require a stock to have at least 10 days of valid return and volume data to compute the ratios in the estimation month. We obtain the data on stock price, return, trading volume, and shares outstanding from the Center for Research in Security Prices (CRSP) daily file and construct monthly Amihud measures. We follow the literature (e.g., Brennan, Huh, and Subrahmanyam 2013) and match the Amihud measures of month $t-2$ to stock returns in month t, and the period of our return analysis is from January 1964 to December 2012. Our main analyses use the monthly measure because it reflects more recent information, and we report the robustness tests using the annual measure.

Panel A of Table 1 presents summary statistics of the Amihud measure and its various components for the 1,197,252 firm-months in our sample, as well as firm size and book-to-market ratio. Firm size is the market capitalization at the end of the previous year. Book-to-market ratio is the ratio of the book value of equity to the market value of equity, where the book value of equity is defined as stockholders' equity plus balance-sheet deferred taxes and investment tax credit, minus the book value of preferred stock.[10] Panel A shows that the trading

[8] A firm-month is dropped from the sample if the firm's stock is traded in a non-NYSE/AMEX exchange on any day of the calendar year of the month.

[9] We nevertheless conduct robustness tests using the NASDAQ sample and report the results in Section 2.3.

[10] Balance-sheet deferred taxes is the Compustat item TXDB, and investment tax credit is item ITCB. We use redemption value (PSTKRV), liquidation value (PSTKL), or par value (PSTK), in that order, for the book value of preferred stock. Stockholders' equity is reported by Moody's (see Davis, Fama, and French 2000) or by

Table 1
Summary statistics and correlations

A. Summary statistics

	Mean	STD	Q10	Q25	Q50	Q75	Q90
A	3.133	14.978	0.001	0.008	0.101	0.861	4.908
AT	35.87	70.75	2.46	5.68	14.46	35.87	80.97
A_C	116.35	329.34	0.07	0.62	8.29	72.29	302.82
AT_C	2,427.22	3,451.93	180.74	423.06	1,169.49	2,989.59	5,993.75
|Ret|	0.020	0.014	0.008	0.011	0.016	0.024	0.035
ME ($M)	2,303.2	11,717.0	10.8	35.1	179.9	964.3	3,712.9
B/M	0.987	0.976	0.249	0.438	0.747	1.218	1.889

B. Correlations among Amihud measures

	A	AT	A_C	AT_C	|Ret|
A	1.000				
AT	0.691	1.000			
A_C	0.899	0.685	1.000		
AT_C	0.312	0.746	0.443	1.000	
|Ret|	0.489	0.347	0.394	−0.040	1.000

Panel A presents summary statistics of the main variables that are constructed monthly from November 1963 to October 2012 for the 1,197,252 firm-months in our sample. Our sample contains ordinary common shares (share codes 10 or 11) listed in NYSE or AMEX. A is the original Amihud (2002) measure, defined as the daily ratio of absolute return to dollar trading volume, averaged across all days in a month. AT is the turnover-based Amihud (2002) measure, defined as the monthly average of the daily ratio of absolute return to turnover, where turnover is daily share volume divided by the shares outstanding. A_C and AT_C are constructed as A and AT, respectively, but the numerators of the ratios are 1 instead of the absolute return. |Ret|is the monthly average of daily absolute return. The Amihud measures, as well as |Ret|, are winsorized at the 1 and 99 percentage points in each cross-section. ME for a firm is the firm's market capitalization at the end of the previous year (in millions of dollars). B/M is the book-to-market ratio calculated as a firm's book value divided by the firm's market capitalization. The B/M ratio is winsorized at the 0.5% and 99.5% level in each cross-section. To ease reading, we multiply A and A_C by 10^6. Panel B presents the time-series averages of the cross-sectional correlation coefficients among the various versions of the Amihud measure and |Ret|. We first calculate cross-sectional correlation coefficients among the variables for each month, and then report the time-series averages of the cross-sectional correlation coefficients.

volume component of the Amihud measure is much more volatile than the return component. The standard deviation of A_C is almost three times its mean, but the standard deviation of $|ret|$ is only 70% of the mean. Additionally, the 75th percentile cutoff of A_C is over 100 times its 25th percentile cutoff, but the 75th percentile cutoff of $|ret|$ is only twice its 25th percentile cutoff. This contrast is also true for the turnover-based Amihud measure. These results suggest that the variation of the trading volume component can account for the majority of the variation in the Amihud measure.

Panel B of Table 1 presents correlations among the various versions of the Amihud measure. We first calculate cross-sectional correlation coefficients among the variables in each month and then report the time-series averages. The Amihud measures are highly correlated with their "constant" measures

---

Compustat (SEQ). If neither is available, we then use the book value of common equity (CEQ) plus the book value of preferred stock. If common equity is not available, stockholders' equity is then defined as the book value of assets (AT) minus total liabilities (LT). We use the book value of the fiscal year ending in calendar year y and market value at the end of year y to calculate book-to-market ratio and match it to stock returns in the one-year period from July of y+1 to June of year y+2. We winsorize the book-to-market ratio in each month at the 0.5% and 99.5% level to reduce the influences of data error and extreme observations.

constructed with only the trading volume components. The correlations are 0.90 between A and A_C and 0.75 between AT and AT_C. These results confirm that the trading volume component alone accounts for a vast majority of the variations in the Amihud measures.

2. Does the Trading Volume Component Explain the Pricing of the Amihud Measure?

We test whether the pricing of the Amihud measure is attributable to its association with trading volume using both sorting and regression analyses.

2.1 Sorting analysis

We sort stocks at the beginning of month t from 1964 to 2012 into quintiles based on their monthly Amihud measures of month $t-2$. We then calculate the equal-weighted portfolio returns each month, and report their time-series averages. The return spreads between the top and bottom quintiles are also reported with the associated t-statistics calculated using Newey and West (1987) standard errors with six lags. We report both raw returns and four-factor alphas calculated using the three Fama-French factors (*MKT*, *SMB*, and *HML*) and the momentum factor (*UMD*).

Panel A of Table 2 presents the sorting analysis for the Amihud (2002) measure (A). The raw return is increasing in the A measure, with the spread between the extreme quintiles being 0.56% per month. This spread is not only economically significant but also statistically significant (t-stat 2.36). The spread in four-factor alpha is 0.35% (t-stat 2.31) per month, which translates to an annual profit of 4.28%. These results are consistent with the regression analyses that the Amihud (2002) measure is strongly related to expected return. When we sort stocks on the "constant" measure, A_C, the return spread is very similar to that of the Amihud measure. The spread is 0.61% per month in raw return and 0.44% in four-factor alpha, both statistically significant. Therefore, excluding the absolute-return component has no impact on the pricing of the Amihud measure.

Next, we use a residual approach to examine whether the A measure is still priced after controlling for the A_C measure. We estimate monthly cross-sectional regressions of the A measure on A_C, and obtain the residuals as the residual A measure. The residual measure therefore represents the variation in the Amihud (2002) measure that is not due to A_C. We sort stocks based on the residual measure, and the results show that a higher residual Amihud measure does not lead to higher expected return. The return spread between the top and the bottom quintiles of the residual measure is insignificantly negative in both raw return (-0.17% , t-stat -1.05) and four-factor alpha (-0.16%, t-stat -0.96).

We further examine AT, the turnover-based Amihud measure, in a similar fashion. Panel B of Table 2 shows that AT has a significantly positive relation

Table 2
Monthly stock returns of portfolios sorted on Amihud measures

	Portfolios sorted on Amihud measures						
	Low	2	3	4	High	H – L	t-stat
A. Sorted on original Amihud measures							
Sorted on A							
Raw return	0.96	1.16	1.23	1.29	1.53	0.56**	(2.36)
Four-factor alpha	−0.03	0.04	0.03	0.10	0.32	0.35**	(2.31)
Sorted on A_C							
Raw return	0.96	1.13	1.23	1.28	1.57	0.61***	(2.95)
Four-factor alpha	−0.05	−0.01	0.03	0.11	0.39	0.44**	(3.20)
Sorted on res. A measure							
Raw return	1.39	1.29	1.19	1.10	1.21	−0.17	(−1.05)
Four-factor alpha	0.25	0.14	0.04	−0.05	0.09	−0.16	(−0.96)
B. Sorted on turnover-based Amihud measures							
Sorted on AT							
Raw return	1.02	1.18	1.25	1.31	1.41	0.39***	(2.64)
Four-factor alpha	−0.19	0.04	0.13	0.19	0.30	0.49***	(3.65)
Sorted on AT_C							
Raw return	1.00	1.26	1.23	1.36	1.33	0.33**	(2.39)
Four-factor alpha	−0.26	0.07	0.10	0.26	0.30	0.55***	(4.47)
Sorted on res. AT measure							
Raw return	1.17	1.18	1.23	1.25	1.35	0.18	(0.74)
Four-factor alpha	0.19	0.06	0.03	0.03	0.16	−0.03	(−0.21)

Panel A presents monthly returns (in percentage) of portfolios sorted on the Amihud measures. A is the monthly Amihud (2002) measure, defined as the daily ratio of absolute return to dollar trading volume, averaged across all days in a month. At the beginning of each month t from 1964 to 2012, stocks are sorted into quintile portfolios according to the A measures of month $t-2$. We then calculate monthly equal-weighted portfolio returns for the quintile portfolios and report time-series average portfolio returns or four-factor alphas, where the four-factor alpha is constructed using the three Fama-French factors and the momentum factor (UMD). The differences between the top and bottom quintiles are also reported with associated t-statistics. We then repeat the sorting for the A_C measure and the residual A measure, where A_C is constructed as A, but the numerator of the ratio is 1 instead of absolute return, and the residual A measure is the residual from the monthly cross-sectional regression of the A measure on the A_C measure. Panel B is similar to panel A, except that we sort stocks based on AT, AT_C, and residual AT, where AT is the turnover-based Amihud (2002) measure, defined as the monthly average of the daily ratio of absolute return to turnover. AT_C is constructed as AT but the numerator of the ratio is 1 instead of absolute return, and the residual AT measure is the residual from the monthly cross-sectional regression of the AT measure on the AT_C measure. The t-statistics (in parentheses) are calculated using Newey-West robust standard errors with six lags.

with expected stock return, and the constant measure AT_C is priced similarly to the AT measure. We then construct a residual AT measure as residuals from monthly cross-sectional regressions of AT on AT_C. When we sort stocks on the residual AT measure, the return spread becomes insignificantly negative (−0.03%, t-stat −0.21).

We also make use of factor returns to examine whether the pricing of the Amihud measure is explained by its trading volume component. This approach is in the same spirit as the approach of using the *SMB* factor, for example, to examine if the abnormal return of a portfolio can be attributed to the size factor. For each month from 1964 to 2012, we sort stocks into terciles according to the “constant” measure A_C of month $t-2$ and then calculate the monthly factor return IML^{A_C} as the equal-weighted return of the top A_C tercile minus that of the bottom A_C tercile.[11] We then repeat the sorting analysis of the Amihud (2002) measure in Table 2, but examine the one-factor alpha calculated using the IML^{A_C} factor, and the five-factor alpha calculated using the IML^{A_C} factor, the three Fama-French factors, and the momentum factor. The results, reported in Table A.1 of the Internet Appendix, show that the positive return premium of the Amihud (2002) measure disappears after controlling for the IML^{A_C} factor. The results are similar when we examine the turnover-based Amihud measure (AT).

2.2 Regression analysis

We further estimate multiple Fama and MacBeth (1973) regressions to examine the pricing of the Amihud (2002) measure. Our first set of the regressions are similar to Brennan, Huh, and Subrahmanyam (2013) by decomposing the Amihud (2002) measure into various components. Brennan, Huh, and Subrahmanyam (2013) decompose the Amihud measure into the turnover-based Amihud measure and firm size (market capitalization) like in Equation (5). They examine these two metrics with regressions of stock returns, and suggest that removing the impact of firm size clarifies the effect of the Amihud measure on stock return. Florackis, Gregoriou, and Kostakis (2011) also use the turnover-based Amihud measure to examine UK stocks. Because our focus is trading volume, we decompose the Amihud (2002) measure into the trading volume component (the A_C measure) and the absolute return component like in Equation (6), and further into the turnover component (the AT_C measure), the absolute return component, and the firm size component like in Equation (7):

$$\ln(A)=\ln\left(\frac{|ret|}{Dvol}\right)=\ln\left(\frac{|ret|}{TO}\times\frac{1}{S}\right)=\ln(AT)-\ln(S), \tag{5}$$

$$\ln(A)=\ln\left(\frac{|ret|}{Dvol}\right)=\ln(|ret|)+\ln\left(\frac{1}{Dvol}\right)=\ln(|ret|)+\ln(A\_C), \tag{6}$$

$$\ln(A)=\ln\left(\frac{|ret|}{Dvol}\right)=\ln\left(|ret|\times\frac{1}{TO}\times\frac{1}{S}\right)=\ln(|ret|)+\ln(AT\_C)-\ln(S), \tag{7}$$

where S is the average daily market capitalization in the estimation month, and the remaining variables are as previously defined. We compute the natural logs of the monthly averages of various daily components: $|\mathrm{ret}|, A\_C, AT, AT\_C$, and S, and estimate regressions of stock returns on these components. We follow Brennan, Chordia, and Subrahmanyam (1998) and use the Fama-French three-factor adjusted return (henceforth FF3-adjusted return) as dependent variable of the return regressions. FF3-adjusted return of firm i in month t is defined as

$$r_{it}^{ff3}=(r_{it}-r_{ft})-(\hat{\beta}_{it}^{MKT}\times MKT_t+\hat{\beta}_{it}^{SMB}\times SMB_t+\hat{\beta}_{it}^{HML}\times HML_t), \tag{8}$$

[11] The results are similar when we construct factor returns by sorting stocks into two or four portfolios instead of three portfolios.

where $\hat{\beta}_{it}^{MKT}$, $\hat{\beta}_{it}^{SMB}$, and $\hat{\beta}_{it}^{HML}$ are estimated for each firm using the monthly excess returns and the three Fama-French factors in the previous 60-month window from $t-60$ to $t-1$.[12] We perform cross-sectional regressions and report the time-series averages of coefficients and the associated t-statistics using the Newey-West (1987) standard errors with six lags. We also include the usual control variables, such as size, book-to-market ratio, and past stock returns, that control for momentum and short-term price reversal. When a regression includes the size component of the Amihud measure (S), we drop the control variable of firm size (market capitalization at the end of previous year). The results are reported in panel A of Table 3.

Model (1) of panel A revisits the pricing of the Amihud measure by regressing return on $\ln(A)$, where the coefficient of $\ln(A)$ is significantly positive, confirming a positive return premium of the Amihud measure. Model (2) regresses return on $\ln(AT)$ and $\ln(S)$ as the decomposition in Equation (5). The results are consistent with Brennan, Huh, and Subrahmanyam (2013) in that the turnover-based Amihud measure is priced. Model (3) decomposes $\ln(A)$ into the volume component ($\ln(A\_C)$) and the absolute return component ($\ln(|ret|)$) like in Equation (6). The coefficient of $\ln(A\_C)$ is positive and significant at the 0.01 level but the coefficient of $\ln(|ret|)$ is significantly negative. Model (4) presents the full decomposition of the Amihud measure like in Equation (7). Although the coefficient of $\ln(AT\_C)$ is significantly positive at the 0.01 level, the coefficient of $\ln(S)$ is significantly negative, and the coefficient of $\ln(|ret|)$ is negative and marginally significant. Overall, panel A shows that the trading volume component of the Amihud measure is positively related to expected return but the absolute return component is not.

In panels B and C, we formally test whether the pricing of the Amihud measure is due to its association with trading volume. Model (1) in panel B is the same as Model (1) in panel A to facilitate comparison with other models. The estimated coefficient of 0.119 for *ln(A)* implies that one standard deviation increase in *ln(A)* (2.69 in our sample period) is associated with a monthly return of 0.32%, in line with the 0.35% alpha spread in the sorting analysis (Table 2). In Model (2), the coefficient of the "constant" Amihud measure ($\ln(A\_C)$) is also significantly positive, indicating that this measure also leads to a return premium. With an estimated coefficient of 0.183 for *ln(A_C)* in Model 2, a one-standard-deviation change (2.53) in *ln(A_C)* leads to an increase in monthly return by 0.46%. In Model (3), we regress return on the residual *ln(A)* measure, which is the residual from the monthly cross-sectional regressions of *ln(A)* on *ln(A_C)*. The coefficient of residual *ln(A)* is significantly negative. Model (4) includes both components of the Amihud measure, *ln(A_C)* and residual *ln(A)*. The coefficient of *ln(A_C)* continues to be significantly positive, and that on residual *ln(A)* remains significantly negative.

[12] We require at least 24 observations in the estimation of factor loadings. We thank Professor Kenneth French for making the data for factor returns available.

Table 3
Monthly Fama-MacBeth regressions of stock returns: Decomposing the Amihud measure

Dependent variable: FF3-adjusted returns

A. Decomposing the Amihud measure

	(1)	(2)	(3)	(4)
ln(A)	0.119***			
	(3.01)			
ln(AT)		0.254***		
		(4.99)		
ln(A_C)			0.165***	
			(3.92)	
ln(AT_C)				0.202***
				(4.57)
ln|Ret|			−0.319**	−0.265*
			(−2.13)	(−1.67)
ln(S)		0.012		−0.052**
		(0.53)		(−2.11)
ln(ME)	0.080**		0.091**	
	(2.04)		(2.01)	
B/M	0.044	0.030	0.011	0.004
	(1.04)	(0.72)	(0.26)	(0.10)
Ret[-12,-2]	0.430**	0.355*	0.320*	0.330*
	(2.05)	(1.67)	(1.74)	(1.68)
Ret[-1]	−6.755***	−6.942***	−7.242***	−7.232***
	(−12.88)	(−13.24)	(−13.80)	(−13.78)
Adj. R^2	0.031	0.031	0.037	0.037
Ave. # obs	1,775	1,775	1,775	1,775
# months	588	588	588	588

B. Regressions on original Amihud measures

	(1)	(2)	(3)	(4)	(5)
ln(A)	0.119***				
	(3.01)				
ln(A_C)		0.183***		0.130***	0.120***
		(4.79)		(3.31)	(3.11)
Res. ln(A)			−0.383***	−0.248***	−0.303***
			(−5.07)	(−3.14)	(−4.65)
Idio. Vol.					−2.903
					(−0.57)
ln(ME)	0.080**	0.144***	−0.067***	0.069*	0.062*
	(2.04)	(3.14)	(−2.48)	(1.82)	(1.74)
B/M	0.044	0.034	0.028	0.026	0.007
	(1.04)	(0.81)	(0.66)	(0.61)	(0.16)
Ret[-12,-2]	0.430**	0.542***	0.372*	0.423**	0.408**
	(2.05)	(2.63)	(1.88)	(2.09)	(2.02)
Ret[-1]	−6.755***	−6.716***	−6.910***	−6.915***	−7.066***
	(−12.88)	(−12.83)	(−13.27)	(−13.23)	(−13.42)
Adj. R^2	0.031	0.032	0.033	0.035	0.040
Ave. # obs	1,775	1,775	1,775	1,775	1,775
# months	588	588	588	588	588

(continued)

In Model (5), we further control for idiosyncratic return volatility, defined as standard deviation of residuals from regressions of a firm's daily returns on the daily Fama-French three factors in the previous year. We control for return volatility as the absolute return component of the Amihud measure is positively correlated with return volatility, and the idiosyncratic volatility is known to affect future returns (e.g., Ang et al. 2006). Model (5) shows that the

Table 3
Continued

C: Regressions on turnover-based Amihud measures

	(1)	(2)	(3)	(4)	(5)
ln(AT)	0.163***				
	(3.95)				
ln(AT_C)		0.223***		0.198***	0.192***
		(5.77)		(5.10)	(5.12)
Res. ln(AT)			−0.246***	−0.207**	−0.267***
			(−3.02)	(−2.53)	(−3.79)
Idio. Vol.					−2.768
					(−0.54)
ln(ME)	−0.025	−0.027	−0.069***	−0.053**	−0.058**
	(−0.99)	(−0.91)	(2.96)	(−2.15)	(−2.51)
B/M	0.041	0.027	0.044	0.025	0.006
	(0.97)	(0.65)	(1.04)	(0.58)	(0.13)
Ret[-12,-2]	0.453**	0.505**	0.334*	0.429**	0.415**
	(2.18)	(2.49)	(1.67)	(2.13)	(2.06)
Ret[-1]	−6.725***	−6.734***	−6.919***	−6.897***	−7.051***
	(−12.86)	(−12.91)	(−13.28)	(−13.23)	(−13.42)
Adj. R^2	0.031	0.032	0.032	0.035	0.040
Ave. # obs	1,775	1,775	1,775	1,775	1,775
# months	588	588	588	588	588

This table presents monthly Fama-MacBeth regressions of stock returns on the components of the Amihud (2002) measure from 1964 to 2012. The dependent variable is the monthly FF3-adjusted return of month t, calculated based on the Fama-French three-factor model, where the factor loadings are estimated in the preceding 60 months. In panel A, the independent variables include the natural logs of the Amihud measure and its components in month $t-2$. A is the original monthly Amihud measure, and A_C is defined as A but the numerator of the ratio is 1 instead of absolute return. AT is the turnover-based Amihud (2002) measure, and AT_C is constructed as AT but the numerator of the ratio is 1 instead of absolute return. $|Ret|$ is the average of daily absolute return over the estimation month. S is the monthly average of daily market capitalization over the estimation month. We also control for a number of firm characteristics. ME is a firm's market capitalization at the end of the previous year (in millions of dollars). B/M is the book-to-market ratio calculated as a firm's book value divided by the firm's market capitalization. For the regression of month t, $Ret[-12,-2]$ is the cumulative stock return from month $t-12$ to month $t-2$, and $Ret[-1]$ is the stock return of month $t-1$. We estimate a cross-sectional regression in each month and then report the time-series means and t-statistics (in parentheses). We also report the time-series averages of the number of observations and adjusted R^2 of the cross-sectional regressions. All the regressions include a constant, which is not reported for brevity. In panel B, *Res.* $ln(A)$ is the residual from the monthly cross-sectional regression of $ln(A)$ on $ln(A\_C)$, and we also control for idiosyncratic return volatility (*Idio. vol.*), defined as standard deviation of residuals from regressions of a firm's daily returns on the daily Fama-French three factors in the previous year. In panel C, *Res.* $ln(AT)$ is the residual from the monthly cross-sectional regression of $ln(AT)$ on $ln(AT\_C)$. t-statistics are calculated using Newey-West robust standard errors with six lags. ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

coefficients of both $ln(A\_C)$ and residual $ln(A)$ are unaffected by the control of idiosyncratic return volatility.

We observe a significantly positive coefficient of firm size, as found by Brennan, Huh, and Subrahmanyam (2013). This result does not mean that larger firms have higher expected returns, because firm size is also a part of the Amihud measure. To illustrate this point, the coefficient of firm size is no longer significantly positive in panel C, which examines the turnover-based Amihud measure that excludes the firm-size component.

In panel C, the coefficients of *ln(AT)* and *ln(AT_C)* are significantly positive when these measures enter the return regressions separately. The estimated coefficient of 0.163 for *ln(AT)* implies that a one-standard-deviation increase in *ln(AT)* (1.10) is associated with a monthly return premium of 0.18%. A one-standard-deviation change (1.09) in *ln(AT_C)* leads to an increase in monthly return by 0.24%. Not surprisingly, these return premiums are lower than those in panel A because the size effect is removed in the turnover versions of the Amihud measures. When *ln(AT_C)* and the residual *ln(AT)* are included in the regression, the coefficient of ln(*AT_C*) is positive and significant at the 0.01 level, but on the residual *ln(AT)* it is significantly negative.

2.3 Robustness tests

We conduct a battery of robustness tests as discussed below. For brevity, we report these results and discuss the details in Section A.1 of the Internet Appendix.

Our first robustness test uses annual Amihud measures instead of monthly measures, where we follow Amihud (2002) and construct annual Amihud measures. Our second robustness test examines the pricing of the Amihud measure for NASDAQ stocks, because our main analyses use NYSE- and AMEX-listed stocks. Our third robustness test repeats the regression analysis separately for the two equal subperiods of 1964–1988 and of 1989–2012. In our fourth robustness test, we repeat the regression analysis using standardized ranks of the independent variables to align the scales of the measures and further control for outliers. Specifically, in each cross-section, we convert the independent variables into uniform distributions between 0 and 1, where 0 corresponds to the lowest value and 1 the highest value. In our fifth robustness test, we consider two "intermediate version" of Amihud measures, *A_C2* and *AT_C2*, where we first calculate daily ratio of absolute return to average daily dollar trading volume over the month, and then average the daily ratios across all days in a month. The results of these robustness tests are consistent with our main analyses in that the pricing of the Amihud measure is explained by its trading volume component.

Our main analyses use the "constant" measures to retain the volume component of the Amihud measures. Because the "constant" measures are the monthly averages of daily reciprocal of dollar trading volume or turnover, they could have distributions and properties different from the dollar trading volume and turnover themselves. We therefore repeat the regression analyses using the monthly average of daily dollar trading volume (*ln(VOLUME*)) or turnover (*ln*(*TO*)) directly, where the results show that our findings hold when we directly examine dollar trading volume or turnover.

2.4 "Half" and "directional" Amihud measures

Brennan, Huh, and Subrahmanyam (2013) propose two "half" Amihud measures constructed using the return-to-turnover ratio on the positive and negative return days separately. They find that, although both "half" measures are associated with a return premium when examined separately, in the multiple return regression framework only the down-day half measure commands a return premium. We therefore examine if the pricing of the "half" Amihud measures is also due to their trading volume component.

The down-day and up-day "half" Amihud measures, *AN* and *AP*, are constructed using the return-to-volume ratios on the negative and positive return days, respectively:

$$AN_{it} = \frac{1}{D_{it}} \sum_{d=1}^{Dit} \frac{-\min[r_{id}, 0]}{Dvol_{id}}, \tag{9}$$

$$AP_{it} = \frac{1}{D_{it}} \sum_{d=1}^{Dit} \frac{\max[r_{id}, 0]}{Dvol_{id}}, \tag{10}$$

where the r_{id} and $Dvol_{id}$ are daily return and daily dollar volume for stock i on day d; D_{it} is the number of days with available ratio in month t.[13] We construct the "constant" measures *AN_C* and *AP_C* corresponding to *AN* and *AP* by replacing the numerator of the daily ratio with a constant one when the ratio is nonzero. We also construct "half" measures corresponding to the turnover-based Amihud measure, *ATN* and *ATP,* where the denominator is daily turnover instead of dollar trading volume.

Panel A of Table 4 presents regression analyses for the *AN* and *AP* measures. Consistent with Brennan, Huh, and Subrahmanyam (2013), both *AN* and *AP* are associated with a return premium when examined separately. More importantly, their constant measures, *AN_C* and *AP_C*, are priced similarly to the half Amihud measures, but the residual half measures are not priced. Panel B of Table 4 examines the *ATN* and *ATP* measures, and the results are similar. These results suggest that the pricing of the "half" Amihud measures is also due to their trading volume component.

Brennan, Huh, and Subrahmanyam (2013) also suggest two "directional" turnover-based Amihud measures based on buy and sell volumes. We follow their approach and separate the trading volume into buy and sell volumes using the Lee and Ready algorithm, and construct *ATNS* and *ATPB*, where *ATNS* (*ATPB*) is constructed similarly to *ATN (ATP)*, but the denominator of the daily ratio is daily sell (buy) turnover. We also construct the constant versions of these two directional measures (labeled *ATPB_C* and *ATNS_C*). Panel C of Table 4 repeats the regression analysis for these four measures, and the results indicate that the pricing of the two directional turnover-based Amihud measures (*ATPB* and *ATNS*) is also explained by their trading volume component (*ATPB_C* and *ATNS_C*).

We further include the pairs of half or directional Amihud measures simultaneously in the return regressions. In panel D of Table 4, Models (1) to (3) show that the coefficients of the down-day "half" measures or sell-volume

[13] We require a stock to have at least 10 days with valid returns and volume data in the estimation month to compute the ratios, and at least two positive return days and two negative return days in the estimation month.

Table 4
Monthly Fama-MacBeth regressions of stock returns: "Half" and "directional" Amihud measures

Dependent variable: FF3-adjusted return

A. Regressions on the negative/positive Amihud measures: Monthly measures

	(1)	(2)	(3)		(4)	(5)	(6)
ln(AN)	0.121***			ln(AP)	0.087**		
	(3.55)				(2.35)		
ln(AN_C)		0.172***	0.123***	ln(AP_C)		0.152***	0.087**
		(5.17)	(3.42)			(4.25)	(2.29)
Res. ln(AN)			−0.220**	Res. ln(AP)			−0.276***
			(−2.33)				(−3.28)
Controls	Yes	Yes	Yes	Controls	Yes	Yes	Yes

B. Regressions on the negative/positive turnover-based Amihud measures

	(1)	(2)	(3)		(4)	(5)	(6)
ln(ATN)	0.219***			ln(ATP)	0.264***		
	(4.36)				(5.33)		
ln(ATN_C)		0.239***	0.219***	ln(ATP_C)		0.258***	0.253***
		(5.45)	(4.70)			(6.05)	(5.75)
Res. ln(ATN)			−0.137	Res. ln(ATP)			−0.034
			(−0.99)				(−0.25)
Controls	Yes	Yes	Yes	Controls	Yes	Yes	Yes

C. Regressions on the directional turnover-based Amihud measures

	(1)	(2)	(3)		(4)	(5)	(6)
ln(ATNS)	0.088***			ln(ATPB)	0.053		
	(2.55)				(1.52)		
ln(ATNS_C)		0.157***	0.136***	ln(ATPB_C)		0.117***	0.100***
		(4.41)	(4.14)			(3.18)	(2.97)
Res. ln(ATNS)			−0.226*	Res. ln(ATPB)			−0.248**
			(−1.79)				(−2.20)
Controls	Yes	Yes	Yes	Controls	Yes	Yes	Yes

(continued)

"directional" measures remain significantly positive and those on the up-day or buy-volume measures are insignificant and close to zero. This result verifies the finding in Brennan, Huh, and Subrahmanyam (2013) that the down-day half measure is priced, but not the up-day half measure, when both are included in the same regression. Models (4) to (6) re-estimate the regressions but use the constant "half" or "directional" measures, and the results show that the constant down-day measures are priced but the constant up-day measures are not. These results suggest that the observed asymmetric relations between the half or directional Amihud measures and expected return also result from their trading volume component.

3. Does Price Impact or the Bid-Ask Spread Explain the Pricing of the Amihud Measure?

Our findings so far show that the pricing of the Amihud measure is explained by its association with trading volume. A natural question, therefore, is whether the pricing of the trading volume component of the Amihud measure is due to the compensation for price impact. Or, is the volume component priced due to

Table 4
Continued

D. Regressions on the negative and positive Amihud measures: Horse race

	(1)	(2)	(3)		(4)	(5)	(6)
ln(AN)	0.143***			ln(AN_C)	0.134***		
	(3.54)				(4.49)		
ln(AP)	−0.021			ln(AP_C)	0.052		
	(−0.46)				(1.33)		
ln(ATN)		0.166***		ln(ATN_C)		0.164***	
		(4.18)				(5.60)	
ln(ATP)		0.000		ln(ATP_C)		0.060	
		(0.01)				(1.55)	
ln(ATNS)			0.107**	ln(ATNS_C)			0.136***
			(2.21)				(4.14)
ln(ATPB)			−0.016	ln(ATPB_C)			0.036
			(−0.33)				(0.91)
Controls	Yes	Yes	Yes	Controls	Yes	Yes	Yes

This table presents the estimation results of monthly Fama-MacBeth regressions of stock returns on the half and directional Amihud measures from 1964 to 2012 (1983 to 2012 when buy and sell volumes are used). The dependent variable is the monthly FF3-adjusted return. FF3-adjusted return of month t is calculated based on the Fama-French three-factor model, where the factor loadings are estimated over the preceding 60 months $[t-60, t-1]$ with at least 24 observations for each firm-level time-series regression. The independent variables are monthly Amihud measures of month $t-2$. In panel A, the independent variables are half Amihud measures. *ln*(*AN*) is the natural log of the monthly half Amihud measure for negative return days (*AN*), which is constructed as *A* but the absolute return-to-volume ratio is nonzero for only the negative return days. *ln*(*AN_C*) is the natural log of *AN_C*, which is constructed as *AN* but the numerator of the ratio is 1 instead of absolute return. *Res. ln*(*AN*) is the residual from the monthly cross-sectional regression of *ln*(*AN*) on *ln*(*AN_C*). *AP* is the monthly half Amihud measure for positive return days, and *AP_C* is the constant version of the *AP* measure. *Res. ln*(*AP*) is the residual from the monthly cross-sectional regression of *ln*(*AP*) on *ln*(*AP_C*). Panel B is similar to panel A, except that independent variables are the half turnover-based Amihud measures. *ATN* and *ATN_C* (*ATP* and *ATP_C*) are constructed as *AN* and *AN_C* (*AP* ad *AP_C*), except that the denominator of the daily ratio is turnover instead of dollar trading volume. Panel C is similar to panel A, except that independent variables are the directional turnover-based Amihud measures. *ATNS* and *ATNS_C* are constructed as *ATN* and *ATN_C*, except that the denominator of the daily ratio is sell turnover (sell volume divided by total shares outstanding). *ATPB* and *ATPB_C* are constructed as *ATP* and *ATP_C*, except that the denominator of the daily ratio is buy turnover (buy volume divided by total shares outstanding). Panel D includes the pairs of half measures or directional measures in the same regressions. The regressions also control for size, book-to-market ratio, momentum, reversal, and an intercept, but for brevity their coefficients are not reported. t-statistics are calculated using Newey-West robust standard errors with six lags. ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

its association with bid-ask spread, the commonly used liquidity benchmark? In this section, we address these questions by controlling for the high-frequency price impact and spread benchmarks in regressions of stock returns on the Amihud measures.

3.1 The construction of high-frequency liquidity benchmarks

We first examine the high-frequency price impact measure, λ, a cost-per-dollar-volume measure widely used in the existing literature (Hasbrouck 2009; Goyenko, Holden, and Trzcinka 2009). Previous studies construct this high-frequency price impact benchmark using the intra-day high-frequency trading data and examine how well the low-frequency liquidity proxies capture price impact. We obtain the transaction data for NYSE/AMEX stocks from 1983 to 2012, including the ISSM data from 1983 to 1992 and the TAQ data from 1993 to 2012. We follow the literature to clean the quotes and trades data, and apply a list of filters on quotes data before calculating NBBO as detailed in Appendix A. We also adopt the methodology in Holden and Jacobsen (2014) to match the trade and quote data for the post-2006 period.

We then follow the literature (Hasbrouck 2009; Goyenko, Holden, and Trzcinka 2009) and estimate the price impact benchmark as the slope coefficient λ of the following regression for each firm-month:

$$r_n = \lambda \times SVol_n + u_n, \tag{11}$$

where for the nth 5-minute period, r_n is the 5-minute stock return calculated as the natural log of the price change over the nth period. (We use quote midpoint instead of trade price to calculate the returns).[14] $SVol_n$ is the signed square-root dollar volume of the nth period, and u_n is the error term. We calculate signed square-root dollar volume as $SVol_n = \sum_{k=1}^{K_n} sign_k \times \sqrt{dvol_k}$, where $dvol_k$ is the dollar volume of the kth trade in the nth 5-minute period, K_n is the number of trades in the nth period, and $sign_k$ is the sign of the kth trade assigned according to the Lee and Ready (1991) trading classification method or the tick test.[15]

To corroborate the analysis using the cost-per-dollar-volume λ measure, we also examine a non-volume-based percentage price impact measure (PI), proposed by Goyenko, Holden, and Trzcinka (2009). Unlike the λ measure, which evaluates the price response to trading volume, the PI measure evaluates the permanent price change of a given trade. Specifically, the percentage 5-minute price impact for a trade is defined as the dollar effective spread minus the dollar realized spread, scaled by the prevailing midpoint five minutes after the trade. We then calculate the monthly PI measure as the average PI for all trades in the estimation month.

In addition to the price impact benchmarks, the high-frequency spread measures are also widely used by the existing literature as liquidity benchmarks. We therefore extend the analysis to the three widely used high-frequency spread benchmarks (Goyenko, Holden, and Trzcinka 2009; Fong, Holden, and Trzcinka Forthcoming): (1) percentage quoted spread (QS): defined as the difference between the bid and ask quote, divided by the midpoint; (2) percentage effective spread (ES): defined as $2 \times |P_k - M_k|$, where P_k is the price of the k^{th} trade, and M_k is the prevailing midpoint for the k^{th} trade. We divide the dollar effective spread by the midpoint and obtain the percentage effective spread (ES); and (3) percentage realized spread (RS). We first calculate the dollar realized spread as $2 \times Sign_k \times |P_k - M_{k+5}|$, where M_{k+5} is the prevailing midpoint 5 minutes after the kth trade, and $sign_k$ is the sign of the kth trade assigned according to the Lee and Ready (1991) trading classification

14 For the return calculation, the opening trade of each day is deleted to remove the overnight return impact.

15 We require at least ten observations in the regressions of monthly λ estimation. Some of the monthly λ estimates (1.12%) are negative and dropped from the regressions after taking the logarithm.

method or the tick test. Dividing the dollar realized spread by M_{k+5} yields the percentage realized spread (RS). We calculate the monthly averages of these spread measures. We winsorize all the high-frequency liquidity benchmarks at the 1st and 99th percentage points in each cross-section to control for outliers.

Panel A of Table 5 presents the summary statistics of these high-frequency liquidity benchmarks and Amihud measures, as well as their correlations from 1983–2012. We calculate cross-sectional correlations and then report their time-series averages. The summary statistics of liquidity benchmarks are close to what the literature has documented. For example, λ in our sample (multiplied by 10^6 to ease reading) has a mean of 30.85 with a standard deviation of 86.12, comparable with the mean (31.93) and standard deviation (88.40) of the random sample of NYSE and AMEX stocks used in Hasbrouck (2009).[16] Our effective spread mean is 1.29% and the sample average of effective half-spread in the random sample of Hasbrouck (2009) is 0.65% (equivalent to an effective spread of 1.30%=0.65% $\times$ 2).[17]

Consistent with Hasbrouck (2009) and Goyenko, Holden, and Trzcinka (2009), we find that the Amihud measure (A) has a correlation of 0.74 with price impact (λ), suggesting that the Amihud (2002) measure performs well capturing price impact. When we remove the size component of the Amihud measure, the resultant AT measure has a lower correlation of 0.60 with the price impact measure. When we further focus on the constant component of the Amihud measure, the correlation between the resultant AT_C and λ is 0.35.[18]

3.2 Do high-frequency liquidity benchmarks explain the pricing of the Amihud measure?

To answer the question whether high-frequency liquidity benchmarks explain the pricing of the volume component of the Amihud measure, we estimate Fama-MacBeth regressions of return on the constant Amihud measures and control for λ and other high-frequency liquidity measures. We also include the usual control variables such as size, book-to-market ratio, and past returns that control for momentum and short-term return reversal. Panels B and C of Table 5 report the results of the constant versions of the Amihud measures, as well as the robustness tests using the original Amihud measures.

In panel B of Table 5, model (1) reexamines the pricing of $ln(A\_C)$ in the sample period of 1983–2012, when we have high-frequency data, where the

[16] Joel Hasbrouck estimated the high-frequency price impact measure for a sample of approximately 300 firms each year from 1993 to 2005. For this comparative sample, the correlation between our estimated annual measure and his estimate is 0.97. We thank Joel Hasbrouck for providing his estimates on his Web site.

[17] We obtain the comparison statistics from the sample data available on Professor Hasbrouck's Web site. The statistics reported by Hasbrouck (2009) are for the random sample including NASDAQ stocks as well, for which the effective spread is 1.06 $\times$ 2 = 2.12%, which is higher than that in the comparison sample we use that includes NYSE and AMEX stocks only.

[18] For robustness, we also examine the correlations between the price impact benchmark and the half and directional constant Amihud measures (ATN_C, ATP_C, $ATNS\_C$, and $ATPB\_C$), and the correlations overall are around 0.2.

Table 5
Do high-frequency liquidity benchmarks explain the pricing of the Amihud measures?

A. Summary statistics of high-frequency liquidity benchmarks, 1983–2012

	Mean	STD	P10	P25	P50	P75	P90	λ	PI	QS	ES	RS	A	A_C	ResA	AT	AT_C
λ ($\times 10^6$)	30.85	86.12	0.32	1.09	4.50	20.90	73.34	1.00									
PI (%)	0.54	0.87	0.04	0.09	0.23	0.60	1.34	0.84	1.00								
QS (%)	1.63	2.55	0.09	0.29	0.80	1.80	3.80	0.79	0.81	1.00							
ES (%)	1.29	4.36	0.07	0.21	0.59	1.33	2.92	0.76	0.79	0.96	1.00						
RS (%)	0.81	3.92	0.02	0.06	0.30	0.72	1.74	0.61	0.58	0.90	0.94	1.00					
A	1.986	12.579	0.000	0.002	0.019	0.287	2.053	0.74	0.65	0.76	0.73	0.67	1.00				
A_C	65.85	240.28	0.03	0.15	1.45	21.71	149.0	0.75	0.68	0.76	0.71	0.63	0.90	1.00			
Res. A	0.00	5.38	−1.22	0.01	0.16	0.46	1.16	0.15	0.08	0.19	0.22	0.22	0.43	0.00	1.00		
AT	29.58	74.72	1.72	3.48	8.01	22.57	64.93	0.60	0.59	0.67	0.63	0.57	0.75	0.75	0.17	1.00	
AT_C	1945.3	3511.5	128.4	261.1	643.9	1881.5	5159.6	0.35	0.39	0.39	0.35	0.31	0.40	0.55	−0.22	0.76	1.00
Res. AT	0.00	50.16	−17.81	−5.96	−2.02	0.66	10.70	0.50	0.46	0.56	0.55	0.51	0.67	0.49	0.55	0.64	0.00

B. Regressions of FF3-adjusted returns on Amihud measures and high-frequency liquidity benchmarks, 1983–2012

	(1)	(2)	(3)	(4)	(5)		(6)	(7)	(8)	(9)	(10)
ln(A_C)	0.138***	0.181***	0.169***	0.210***		ln(A)	0.568	0.108**	0.115**	0.194***	
	(3.37)	(3.78)	(3.71)	(3.72)			(1.36)	(2.35)	(2.38)	(3.10)	
ln(λ)		−0.154**		−0.066		ln(λ)		−0.120**		−0.062	
		(−2.44)		(−1.56)				(−2.10)		(−1.41)	
ln(PI)			−0.226***			ln(PI)			−0.222**		
			(−2.57)						(−2.41)		
ln(QS)				−0.180		ln(QS)				−0.131	
				(−0.91)						(−0.68)	
ln(ES)				0.253		ln(ES)				−0.406*	
				(−1.23)						(−1.84)	
ln(RS)				0.056		ln(RS)				0.079	
				(1.06)						(1.53)	
Cost component					0.061 (1.39)	Cost component					0.015 (0.35)
Noncost component					0.220*** (3.73)	Noncost component					0.191*** (3.10)
Controls	Yes	Yes	Yes	Yes	Yes	Controls	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.028	0.030	0.030	0.035	0.031	Adj. R2	0.028	0.029	0.029	0.034	0.030
Ave. # obs	1,692	1,692	1,692	1,692	1,692	Ave. # obs	1,692	1,692	1,692	1,692	1,692

(continued)

Table 5
Continued

C. Regressions of FF3-adjusted returns on turnover-based Amihud measures and high-frequency liquidity Benchmarks, 1983–2012

	(1)	(2)	(3)	(4)	(5)		(6)	(7)	(8)	(9)	(10)
ln(AT_C)	0.187***	0.228***	0.214***	0.244***		ln(AT)	0.108***	0.182***	0.178***	0.254***	
	(4.72)	(4.77)	(4.72)	(4.33)			(2.63)	(3.86)	(3.50)	(3.90)	
ln(λ)		−0.163***		−0.075*		ln(λ)		−0.171***		−0.086*	
		(−2.58)		(−1.75)				(−2.77)		(−1.83)	
ln(PI)			−0.229***			ln(PI)			−0.267***		
			(−2.62)						(−2.80)		
ln(QS)				−0.192		ln(QS)				−0.198	
				(−0.99)						(−1.03)	
ln(ES)				−0.212		ln(ES)				0.367*	
				(−1.05)						(−1.66)	
ln(RS)				0.054		ln(RS)				0.084	
				(1.02)						(1.58)	
Cost component					−0.015	Cost component					−0.312**
					(−0.12)						(−2.28)
Noncost component					0.230***	Noncost component					0.248***
					(4.49)						(3.84)
Controls	Yes	Yes	Yes	Yes	Yes	Controls	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.028	0.030	0.030	0.035	0.030	Adj. R^2	0.027	0.029	0.029	0.034	0.030
Ave. # obs	1,692	1,692	1,692	1,692	1,692	Ave. # obs	1,692	1,692	1,692	1,692	1,692

Panel A reports the summary statistics of the high-frequency liquidity benchmarks and the Amihud measures, together with their correlations, in the sample period of 1983–2012. The price impact measure λ is estimated as the slope coefficient of the monthly regression of five-minute stock returns on signed square root dollar volume in the same time period. We require at least ten valid observations for the regressions. The percentage 5-minute price impact (*PI*) is the dollar effective spread minus the dollar realized spread, scaled by M_{k+5}, the prevailing midpoint five minutes after the trade. The dollar effective spread is $2\cdot|P\downarrow k-M\downarrow k$, where P_k is the price of the kth trade, and M_k is the prevailing midpoint for the kth trade. The dollar realized spread, $2Sign_k\cdot|P_k-M_{k+5}|$, divided by the post-trade quotes midpoint M_{k+5}. M_{k+5} is the prevailing midpoint 5 minutes after the kth trade, and $sign_k$ is the sign of the kth trade assigned according to the Lee and Ready (1991) trading classification method or the tick test. *QS* is the percentage quoted spread. *ES* is the percentage effective spread. *RS* is the percentage realized spread. We calculate the means of these spread measures for each stock-month. To control for outliers, we winsorize the high-frequency liquidity measures at the 1 and 99 percentage points in each cross-section. Table 3 provides definitions of the Amihud measures. To ease reading, we multiply *A* and *A_C* by 10^6, and *PI*, *QS*, *ES*, and *RS* by 10^2. In panels B and C, we report the monthly Fama-Macbeth regressions of stock returns on the Amihud measures after controlling for the high-frequency liquidity benchmarks. The dependent variable is the monthly FF3-adjusted return of month t. The independent variables include the natural logs of the monthly Amihud measures and high-frequency liquidity measures of month $t-2$. We also control for firm characteristics including size (*ln(ME)*), book-to-market ratio (B/M), momentum (Ret[−12,−2]), reversal (Ret[-1]), and an intercept, but do not report them for brevity. We also estimate cost component and noncost component of the Amihud measures. For example, for the A_C measure, we first estimate monthly cross-sectional regression of *ln*(*A_C*), natural log of the constant Amihud measure, on the natural logs of the high-frequency liquidity benchmarks including *ln*(λ), *ln*(*QS*), *ln*(*ES*), and *ln*(*RS*). We then measure the transaction-cost component as the fitted value of the regression, and noncost component as the residual of the regression. We estimate the cost and noncost components for the other Amihud measures similarly. We estimate a cross-sectional regression in each month from March 1983 to December 2012 (358 months) and then report the time-series means and t-statistics (in parentheses). t-statistics are calculated using Newey-West robust standard errors with six lags. ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

coefficient of $ln(A\_C)$ remains significantly positive. In Model (2), we add $ln(\lambda)$, the natural log of the price impact measure λ of the month $t-2$. The coefficient of $ln(\lambda)$ is significantly negative, indicating that the price impact benchmark itself is not positively related to expected return, whereas the coefficient of $ln(A\_C)$ remains positive and statistically significant (t-stat 3.78). Model (3) repeats the regression analysis using the *PI* measure instead of λ, where the coefficient of *PI* is insignificantly negative like the λ measure, but the coefficient of $ln(A\_C)$ remains significantly positive after controlling for the *PI* measure, suggesting that this alternative high-frequency price impact measure cannot explain the pricing of the trading volume component of the Amihud measure.

Model (4) further controls for the natural logs of the high-frequency spread benchmarks by including all three spread benchmarks in addition to the price impact benchmark ($ln(\lambda)$). The coefficient is significantly positive for $ln(A\_C)$ (t-stat 3.72), but insignificantly negative for the liquidity benchmarks. The results indicate that the pricing of the trading volume component of the Amihud measure is not due to its association with the spread.

Finally, we decompose the constant Amihud measure into a component associated with transaction costs and a residual component ("noncost" component), and examine the pricing of the two components separately. We first estimate monthly cross-sectional regressions of the natural log of $ln(A\_C)$ on the natural logs of the high-frequency liquidity benchmarks including $ln(\lambda)$, *ln(QS)*, *ln(ES)*, and *ln(RS)*. We then calculate the transaction-cost component of the Amihud measure as the fitted value of the regression, and noncost component as the residual of the regression. The noncost component, therefore, is the part of the Amihud measure that is orthogonal to the transaction cost benchmarks. Model (5) in panel B shows that the coefficient of the transaction cost component is insignificant (t-stat 1.39), but that on the noncost component is significantly positive (t-stat 3.73). These results reveal that, although the Amihud measure is highly correlated with transaction costs, it is the non-transaction-cost component that drives the pricing of the Amihud measure.

In panel C Models (1) to (5), we repeat the analyses for the constant version of the turnover-based Amihud measure (*AT_C*), and the results are similar to those using the *A_C* measure. We also conduct robust tests using the original Amihud measure (*A*) and turnover-based Amihud measure (*A*T) in Models (6) to (10) of panels B and C, respectively. The results using these measures are similar to those using the constant Amihud measures.

We conduct a number of additional robustness tests. First, an alternative approach to measure price impact is estimating trade-based price impact like in Glosten and Harris (1988). We use the λ measure for our analysis because it is closely related to the Amihud measure, both conceptually and empirically, as is shown by the existing literature. Nevertheless, we use the Glosten and Harris' (1988) approach to estimate price impact (result untabulated for brevity), and our finding holds that price impact does not explain the pricing of the Amihud measures. Second, to avoid the estimation of λ being driven by certain days of the estimation month, we estimate daily λ and then average across the days of the estimation month. Third, Easley, Lopez de Prado, and O'Hara (2012) point out that the Lee and Ready algorithm may be more error-prone in the recent high-frequency trading era. As a result, Brennan, Huh, and Subrahmanyam (2013) exclude the post-2006 period from some of their analyses. We therefore conduct the robustness test by excluding the 2006–2012 period. Fourth, we repeat the analyses without skipping a month between the λ measure and stock return, that is, matching λ of month $t-1$ with return in month t. Fifth, we construct annual λ measure instead of monthly measure, and match monthly stock returns with λ of the previous year. For brevity, we report these tests in Tables A.7 and A.8 of the Internet Appendix. They confirm our finding that the price impact benchmark is neither priced nor explains the pricing of the Amihud measure. Fifth, we also repeat the analysis using the annual Amihud measures and liquidity benchmarks instead of monthly measures in Table A.9 of the Internet Appendix, and find similar results.

4. Is the Pricing of Trading Volume Caused by Liquidity Premium or by Mispricing?

Our results so far show that the pricing of the Amihud measure is explained by its association with trading volume, and that the pricing of the trading volume component is unlikely explained by the compensation for price impact or association with other liquidity benchmarks. Then why is trading volume priced? Is the return premium of trading volume, aka volume premium, a liquidity premium or not?

It is worth noting that the source of volume premium is the subject of debate in a quite large finance literature. On one hand, trading volume is generally considered a (noisy) liquidity proxy, and it is possible that the volume premium is a liquidity premium associated with some aspect of liquidity that is not reflected in the high-frequency liquidity benchmarks that we examined. For example, Collin-Dufresne and Fos (2015) and Easley, Lopez de Prado, and O'Hara (2016) suggest the standard measures of adverse-section may fail to capture informed trading, because of long-lived information, limit order usage, and selective informed trader participation on high volume and low volume days. On the other hand, a large number of studies have attributed the pricing of trading volume to various nonliquidity factors, most of which are associated with mispricing (investor disagreement, sentiment, investor attention, etc.).

We attempt to distinguish the liquidity and the nonliquidity explanations of the volume premium with a balanced analysis, including two tests on each side of the argument based on the existing literature. This is a challenge as illustrated by the existence of many studies on this topic, and this topic by itself can constitute a stand-alone paper. We acknowledge that none of our tests are perfect but we believe that together they can shed light on the nature of the volume premium. In our tests, we focus on two clean measures of trading volume from which the size effect is removed: (1) *AT_C*, the constant version of the turnover-based Amihud measure; and (2) *Turnover*, the monthly average of daily turnover. Both are constructed using only the turnover of a stock.

4.1 Tests of the liquidity explanation

4.1.1 Seasonality of the liquidity premium. Our first test of the liquidity explanation is motivated by the literature on the seasonality of liquidity premium. Specifically, Eleswarapu and Reinganum (1993) and Hasbrouck (2009) both find that liquidity premium is significant in January but insignificant in non-January months. If the volume premium is a liquidity premium, then we expect it to demonstrate a similar January seasonality.

In Table 6, we estimate firm-level Fama-MacBeth return regressions on the high-frequency liquidity benchmarks and the two measures of turnover individually, controlling for size, book-to-market ratio, momentum, and reversal. Before conducting the analysis for January and non-January separately, we first report the full-sample results in panel A of Table 6. Similar to the results in Table 5, none of the high-frequency benchmarks has a positive significant premium. In contrast, there is a clear turnover premium as shown by the positive coefficient for *ln*(*AT_C*) and negative coefficient for *ln*(*TO*) in the last two columns.

In panels B and C, we repeat the return regressions for January and non-January separately. Consistent with the seasonality of liquidity premium documented by Eleswarapu and Reinganum (1993) and Hasbrouck (2009), we report in panel B that the coefficients of the liquidity benchmarks are significantly positive in January—except for *PI*, which is insignificantly positive—but the coefficients of the liquidity benchmarks are insignificant or significantly negative in non-January months as shown in panel C.

In a stark contrast, the last two columns of panels B and C suggest that there is an opposite seasonality for volume premium. In January, the coefficient of *AT*_C is significantly negative and on turnover is significantly positive, indicating that the volume premium reverses in January. That is, higher turnover stocks outperform low turnover stocks in January. However, in non-January months, the coefficient of *AT_C* is significantly positive and that on turnover is significantly negative. Therefore, we find that the volume premium exhibits an opposite seasonality to that of the liquidity premium.[19]

Although consistent with the earlier literature, the finding that the liquidity benchmarks are priced only in January is puzzling and unexplained by the

[19] For robustness, we also repeat the regression analysis using annual measures and find similar results. These results are reported in Tables A.10 of the Internet Appendix. We also examine the January seasonality of the pricing of the transaction cost and noncost components and present the results in Table A.11 of the Internet Appendix. The non-transaction-cost component is also priced in non-January months, but not in January.

Table 6
Pricing high-frequency liquidity benchmarks and turnover measures, January versus non-January

						Turnover Measures	
	ln(λ)	ln(PI)	ln(QS)	ln(ES)	ln(RS)	ln(AT_C)	ln(TO)
A. Regressions of stock returns on liquidity benchmarks and turnover, all months							
Liq. Benchmark	−0.033	−0.118	−0.252*	−0.246*	−0.123*	0.187***	−0.198***
	(−0.64)	(−1.57)	(−1.92)	(−1.93)	(−1.94)	(4.72)	(−4.40)
ln(ME)	−0.069*	−0.099***	−0.139***	−0.135***	−0.085***	−0.004	−0.024
	(−1.93)	(−3.02)	(−3.38)	(−3.56)	(−3.74)	(−0.10)	(−0.65)
B/M	−0.028	−0.025	−0.021	−0.018	−0.014	−0.035	−0.036
	(−0.51)	(−0.45)	(−0.39)	(−0.34)	(−0.25)	(−0.64)	(−0.65)
Ret[-12,-2]	0.236	0.224	0.057	0.067	0.143	0.349	0.349
	(0.78)	(0.74)	(0.19)	(0.23)	(0.48)	(1.16)	(1.16)
Ret[-1]	−4.955***	−4.970***	−5.092***	−5.115***	−5.003***	−4.939***	−4.968***
	(−8.79)	(−8.87)	(−9.00)	(−9.01)	(−8.89)	(−8.76)	(−8.80)
B. Regressions of stock returns on liquidity benchmarks and turnover, January							
Liq. Benchmark	0.485**	0.565	3.013***	3.053***	1.489***	−0.381**	0.340*
	(2.15)	(1.11)	(3.77)	(4.20)	(4.56)	(−2.33)	(1.81)
ln(ME)	−0.648***	−0.743***	0.332	0.340	−0.352***	−1.125***	−1.077***
	(−4.51)	(−3.47)	(1.27)	(1.47)	(−4.75)	(−6.80)	(−7.40)
B/M	0.219	0.200	0.112	0.105	0.073	0.213	0.196
	(0.81)	(0.77)	(0.43)	(0.41)	(0.30)	(0.81)	(0.76)
Ret[-12,-2]	−2.936***	−2.944***	−2.521***	−2.484***	−2.528***	−2.909***	−2.896***
	(−4.25)	(−4.35)	(−3.56)	(−3.58)	(−3.62)	(−4.15)	(−4.19)
Ret[-1]	−16.197***	−16.100***	−16.503***	−16.375***	−15.978***	−15.862***	−15.999***
	(−9.50)	(−9.80)	(−9.86)	(−10.05)	(−9.97)	(−9.02)	(−9.09)
C. Regressions of stock returns on liquidity benchmarks and turnover, non-January							
Liq. Benchmark	−0.079	−0.178**	−0.540***	−0.537***	−0.265***	0.237***	−0.246***
	(−1.45)	(−2.18)	(−4.07)	(−4.21)	(−4.24)	(6.54)	(−5.94)
ln(ME)	−0.018	−0.043	−0.180***	−0.177***	−0.062***	0.095**	0.069*
	(−0.53)	(−1.30)	(−4.62)	(−4.98)	(−2.70)	(2.47)	(1.80)
B/M	−0.050	−0.045	−0.033	−0.029	−0.021	−0.057	−0.056
	(−0.93)	(−0.83)	(−0.63)	(−0.56)	(−0.40)	(−1.07)	(−1.05)
Ret[-12,-2]	0.516*	0.504*	0.284	0.292	0.379	0.637**	0.635**
	(1.70)	(1.66)	(0.97)	(1.00)	(1.27)	(2.08)	(2.08)
Ret[-1]	−3.964***	−3.988***	−4.087***	−4.122***	−4.035***	−3.976***	−3.995***
	(−7.08)	(−7.15)	(−7.25)	(−7.28)	(−7.15)	(−7.07)	(−7.10)

This table presents Fama-MacBeth regressions of monthly stock returns on the monthly high-frequency liquidity benchmarks and turnover measures in panel A, and separately for January (panel B) and non-January months (panel C) from March 1983 to December 2012. The dependent variable is the monthly FF3-adjusted return of month t. The independent variables include the natural logs of the monthly high-frequency liquidity measures and turnover measures of month $t-2$. The price impact measure λ is estimated as the slope coefficient of the monthly regression of five-minute stock returns on signed square root dollar volume in the same time period. *QS* is the percentage quoted spread. *ES* is the percentage effective spread. *RS* is the percentage realized spread. We calculate the means of these spread measures for each stock-month. *ln*(*AT_C*) is the natural log of constant version of the turnover-based Amihud measure. *ln(TO)* is the natural log of monthly average of the daily turnover. We also control for size (*ln(ME)*), book-to-market ratio (B/M), momentum (*Ret*[*–12,–2*]), reversal (*Ret*[*-1*]), and an intercept but do not report them for brevity. We estimate a cross-sectional regression in each month and then report the time-series means and t-statistics (in parentheses) calculated using Newey-West robust standard errors with six lags. ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

existing theory of liquidity premium. But the opposite seasonality suggests that the return premium of trading volume and that of the known liquidity benchmarks may be shaped by distinct mechanisms.

4.1.2 Subperiods of stock market illiquidity. Investors demand a liquidity premium because it is costly to liquidate illiquid assets, and such liquidation costs will be higher when market illiquidity is higher (Pástor and Stambaugh 2003). Pástor and Stambaugh also suggest that illiquidity is a state variable and investors whose wealth drops during episodes of high market illiquidity will find greater liquidation costs especially unfavorable. As a result, liquidity premium is expected to be larger in the episodes of high market illiquidity.

Therefore, our second test of the liquidity explanation is to examine the relation between the volume premium and market illiquidity. We use Pástor and Stambaugh's (2003) market illiquidity measure for this analysis because it is a widely used proxy for aggregate illiquidity and it coincides well with the cycle of market illiquidity. We obtain the monthly measure of market illiquidity for our sample period 1964–2012 from Professor Pástor's data library, and our test design follows Stambaugh, Yu, and Yuan (2012), who examine the relation between anomalies and market sentiment except that our variable of interest is market illiquidity instead of market sentiment.[20]

Panel A of Table 7 examines the volume premium across high and low market illiquidity periods. To ease reading, we denote the "long" portfolio as the top quintile of *AT_C* or the bottom quintile of turnover, and the "short" portfolio as the bottom quintile of *AT_C* or the top quintile of turnover. "Long-Short" is therefore the volume premium. The period of high (low) market illiquidity contains the months where market illiquidity of month $t-1$ is above (below) the median. The results show that the volume premium is very similar across the period of high market illiquidity and that of low market illiquidity. For example, the return spread of *AT_C* is 0.33% for the low market illiquidity period and 0.32% for the high illiquidity period, with a difference of just 0.02% (t-stat 0.08).

Next, we estimate time-series regressions of portfolio returns to further examine the effect of market illiquidity on the volume premium and control for return factors:

$$R_{it} = a + bIlliq_{t-1} + cMKT_t + dSMB_t + eHML_t + fMOM_t + u_t, \qquad (12)$$

where R_{it} is the return of turnover-based portfolio i of month t, in excess of risk-free rate. $Illiq_{t-1}$ is market illiquidity measure of month $t-1$. MKT_t, SMB_t, HML_t, and MOM_t are the three Fama-French factors and momentum factor of month t. Stambaugh, Yu, and Yuan (2012) use a similar approach to examine the relation between an anomaly and market sentiment, and our model differs in that our main independent variable is market illiquidity rather than investor sentiment.

Panel B of Table 7 presents the coefficient of the market illiquidity (b) in Equation (12). The dependent variables are returns of quintile portfolios of *AT_C* or turnover, as well as the return spread ("Long-Short"). In the regression of return spread, the coefficient of market illiquidity is significantly negative for

---

[20] We thank Professor Luboš Pástor for making this dataset available.

Table 7
Turnover premium across subperiods of market illiquidity

	ln(λ)	ln(PI)	ln(QS)	ln(ES)	ln(RS)	ln(AT_C)	ln(TO)
A. Returns of portfolios sorted on turnover measures: Periods of high and low market illiquidity							
	Short	2	3	4	Long	Long – Short	t-stat
Sorted on AT_C							
High market illiquidity	1.11	1.36	1.31	1.45	1.44	0.33*	(1.67)
Low market illiquidity	0.89	1.16	1.15	1.27	1.21	0.32	(1.30)
High – Low						−0.02	
t-stat						(−0.08)	
Sorted on TO							
High market illiquidity	1.13	1.38	1.37	1.45	1.33	0.20	(1.00)
Low market illiquidity	0.95	1.14	1.17	1.21	0.21	0.26	(1.06)
High – Low						−0.06	
t-stat						(−0.30)	

B. Coefficients of market illiquidity (b) in the FF4 model
$R_{it} = a + bIlliq_{t-1} + cMKT_t + dSMB_t + eHML_t + fMOM_t + u_t$

	Short	2	3	4	Long	Long – Short
Sorted on AT_C						
Coefficient (b)	0.031**	0.017	0.008	−0.005	−0.016	−0.048**
t-stat	(2.18)	(1.51)	(0.80)	(−0.46)	(−1.11)	(−2.38)
Sorted on TO						
Coefficient (b)	0.008	0.009	−0.011	−0.010	−0.030**	−0.038**
t-stat	(0.58)	(0.87)	(−1.16)	(−0.87)	(−2.14)	(−2.05)

This table reports the results on the relation between the volume premium and market illiquidity. We obtain the monthly measure of market illiquidity for our sample period 1964–2012 from Professor Luboš Pástor's data library and define a month t as period of high (low) market illiquidity if the market illiquidity of month $t-1$ is above (below) the median. The turnover measures used include *AT_C,* the constant version of the turnover-based Amihud measure, and *TO,* the monthly average of the daily turnover. We match returns of month t to turnover measures of $t-2$. Panel A reports the monthly returns of portfolios sorted on the turnover measures across high and low market illiquidity periods. The "long" portfolios are the top (bottom) quintile of *AT_C* (turnover), and the "short" portfolios are the bottom (top) quintile of *AT_C* (turnover). "Long-Short" is the monthly volume premium. In panel B, we estimate the following time-series regressions to further control for return factors: $R_{it} = a + bIlliq_{t-1} + cMKT_t + dSMB_t + eHML_t + fMOM_t + u_t$, where R_{it} is the return of turnover-based portfolio i of month t, in excess of risk-free rate. $Illiq_{t-1}$ is market illiquidity measure of month $t-1$. MKT_t, SMB_t, HML_t, and MOM_t are the Fama-French factors and momentum factor of month t. Panel B presents the coefficient of the market illiquidity (b) and its t-statistics.

both *AT_C* and turnover (t-stats −2.38 and −2.05), indicating that the volume premium is *smaller* in episodes of high market illiquidity after controlling for the return factors. This result does not support the liquidity explanation of the volume premium.

4.2 Tests of mispricing explanation

Next, we turn to the mispricing explanation of the volume premium and conduct two tests based on the existing literature.

4.2.1 Subperiods of market sentiment. Our first test is motivated by Stambaugh, Yu, and Yuan (2012), who suggest that market wide investor sentiment can be used to identify mispricing and especially overpricing. Specifically, Miller's (1977) theory suggests that in the presence of short-sale constraints, overpricing can be caused by a group of over-optimistic investors and will be greater in the period of high market sentiment, which is characterized

Table 8
Turnover premium across subperiods of market investor sentiment

A. Returns of portfolios sorted on turnover measures: Periods of high and low market sentiment

	Short	2	3	4	Long	Long – Short	t-stat
Sorted on AT_C							
Low sentiment	1.54	1.63	1.51	1.75	1.70	0.16	(0.65)
High sentiment	0.52	0.92	0.98	1.00	1.04	0.52**	(2.34)
High - Low	−1.02	−0.71	−0.53	−0.75	−0.66	0.36	
t-stat	(−1.64)	(−1.35)	(−1.10)	(−1.51)	(−1.41)	(1.08)	
Sorted on TO							
Low sentiment	1.61	1.65	1.62	1.63	1.63	0.02	(0.06)
High sentiment	0.54	0.89	0.97	1.07	0.99	0.45**	(2.03)
High - Low	−1.07*	−0.75	−0.65	−0.56	−0.64	0.43	
t-stat	(−1.70)	(−1.39)	(−1.31)	(−1.18)	(−1.43)	(1.30)	

B. Coefficients of sentiment (b) in the FF4 model
$R_{it} = a + bSent_{t-1} + cMKT_t + dSMB_t + eHML_t + fMOM_t + u_t$,

	Short	2	3	4	Long	Long – Short
Sorted on AT_C						
Coefficient (b)	−0.25***	−0.17**	−0.12*	−0.04	−0.01	0.24*
t-stat	(−2.67)	(−2.18)	(−1.81)	(−0.57)	(−0.10)	(1.81)
Sorted on TO						
Coefficient (b)	−0.25***	−0.19***	−0.10	−0.03	−0.02	0.23*
t-stat	(−2.67)	(−2.42)	(−1.47)	(−0.53)	(−0.23)	(1.85)

This table reports the results on the relation between the volume premium and market sentiment. The monthly market-wide investor sentiment index (Baker and Wurgler 2006) is obtained from Professor Jeffrey Wurgler's web site for the period from July 1965 to December 2010. Periods of high (low) sentiment are the months with lagged sentiment index above (below) the median. The turnover measures used include *AT_C,* the constant version of the turnover-based Amihud measure, and *TO,* the monthly average of the daily turnover. We match returns of month t to turnover measures of $t-2$. Panel A reports monthly returns of portfolios sorted on the turnover measures across high- and low-sentiment periods. We define the "long" portfolio as the top (bottom) quintile of *AT_C* (turnover), and the "short" portfolio as the bottom (top) quintile of *AT_C* (turnover). "Long-Short" is the monthly volume premium. In panel B, we estimate the following time-series regression: $R_{it} = a + bSent_{t-1} + cMKT_t + dSMB_t + eHML_t + fMOM_t + u_t$, where R_{it} is return of *AT_C* (or turnover) portfolio i of month t, in excess of risk-free rate. $Sent_{t-1}$ is sentiment index of month $t-1$. MKT_t, SMB_t, HML_t, and MOM_t are the Fama-French factors and momentum factor of month t. Panel B presents the coefficient of the sentiment index (b) and its $t-$statistics.

by a larger number of over-optimistic investors. Stambaugh, Yu, and Yuan (2012) therefore hypothesize that an anomaly associated with overpricing will be much stronger following high market sentiment, driven by the more negative return of the short leg of the anomaly.

Following Stambaugh, Yu, and Yuan (2012), we obtain the monthly market-wide investor sentiment index (Baker and Wurgler 2006) from July 1965 to December 2010, and define the period of high (low) sentiment as the months with lagged sentiment index above (below) the median.[21] Panel A of Table 8 reports the sorting analyses across high- and low-sentiment periods. To ease reading, we define the "long" portfolio as the top quintile of *AT_C* or the bottom quintile of turnover, and the "short" portfolio as the bottom quintile of *AT_C* or the top quintile of turnover. "Long-Short" is the return spread (volume premium).

21 We thank Professor Jeff Wurgler for making the sentiment index available on his Web site.

Panel A shows that stock returns are lower following high sentiment than following low sentiment. This general pattern is consistent with Baker and Wurgler (2006) and Stambaugh, Yu, and Yuan (2012), indicating that stocks are likely to be overvalued after a high sentiment period. More importantly, consistent with Stambaugh, Yu, and Yuan's prediction, the return spread of *AT_C* or turnover is much larger following high market sentiment than following low market sentiment, and this difference is driven by the more negative return of the short leg. For example, for the *AT_C* portfolios, the short leg is 1.02% lower following high market sentiment than following low market sentiment, which is much larger than the corresponding 0.66% difference for the long leg.

We further estimate time-series regressions to controls for the return factors:

$$R_{it} = a + bSent_{t-1} + cMKT_t + dSMB_t + eHML_t + fMOM_t + u_t, \qquad (13)$$

where R_{it} is return of *AT_C* or turnover portfolio i of month t, in excess of risk-free rate. $Sent_{t-1}$ is sentiment index of month $t-1$. MKT_t, SMB_t, HML_t, and MOM_t are the Fama-French factors and momentum factor of month t. Panel B of Table 8 presents the coefficient of the sentiment index (b). For both *AT_C* and turnover, the coefficient for the short leg is significantly negative, indicating a negative relation between short leg return and market sentiment. The coefficient for the long leg, on the contrary, is insignificant, suggesting that market sentiment does not have a significant impact on the long leg return. The coefficient of the return spread ("Long-Short") is significantly positive, indicating that the volume premium is stronger following episodes of high market sentiment. These results support the mispricing explanation of the volume premium.

4.2.2 Earnings announcement and non-earnings-announcement periods. Our second test of the mispricing explanation examines the volume premium in the earnings announcement period and non-announcement period separately. This approach is proposed by La Porta et al. (1997), who suggest that an anomaly associated with mispricing will be more pronounced in an earnings announcement period as earnings news helps correct mispricing. A contemporaneous study by Engelberg, McLean, and Pontiff (2016) also uses this approach to examine a strategy that combines 94 anomalies.[22] If the volume premium is due to high volume stocks being overpriced and therefore earning low future returns, then we expect the volume premium to be more pronounced in the earnings announcement period than in the non-earnings-announcement period.

[22] Engelberg, McLean, and Pontiff (2016) combine 94 anomalies into a single strategy instead of separately examining each. They find that the abnormal return of this strategy is much stronger in earnings announcement and news days than other days.

We collect earnings announcement dates from Compustat from 1972 to 2012, because the announcement dates are available from 1972. At the beginning of month t, we examine the subset of sample firms with earnings announcement in month t, and classify these stocks into quintile portfolios according to AT_C or turnover of month $t-2$. We use the full sample ranks of AT_C or turnover to form portfolios in case the full sample distribution differs from that of the announcement sample. Then for each announcing stock in month t, we calculate buy-and-hold abnormal return in the 3-day window $[-1,1]$ surrounding the announcement date (BHAR $[-1,1]$) as the buy-and-hold raw return minus the buy-and-hold value-weighted CRSP return.[23] We also calculate buy-and-hold abnormal returns for the non-announcement days in month t. We then calculate portfolio BHARs every month and report time-series averages.

Panel A of Table 9 shows that BHAR $[-1,1]$ is increasing in AT_C, and the spread between the top and the bottom quintiles is a large 0.63% (t-stat 7.53). We observe very similar results when sorting stocks on the turnover measure. In stark contrast, BHAR for the non-announcement period does not vary much across AT_C or turnover despite the fact that the non-announcement period is much longer than the announcement window. For example, the spread in non-announcement BHAR is an insignificantly negative -0.10% (t-stat -0.73) for AT_C. Panel B further presents Fama-MacBeth regressions of BHARs on AT_C or turnover, with the usual controls of firm size, book-to-market ratio, momentum, and short-term return reversal. In the regression of announcement BHAR, the coefficient is significantly positive for AT_C and significantly negative for turnover. In the regressions of non-announcement BHAR, however, the coefficient of AT_C or turnover becomes insignificant and the sign flips. These results support the sorting analyses that the volume premium is concentrated in the earnings announcement window.[24]

We further investigate analyst forecast errors to directly examine if investors in high volume stocks are over-optimistic. Analyst forecast error for an announcement is the consensus forecast minus actual earnings, scaled by stock price at the end of the previous quarter.[25] A positive forecast error suggests over-optimism, and a negative error suggests investor over-pessimism. If high volume stocks are more likely to be associated with

---

23 If only part of the 3-day earnings announcement window $[-1,1]$ falls in month t, we do not drop the announcement but use the partial earnings announcement return in month t. We conduct robustness tests by dropping these partial announcements, and the results are similar.

24 To address the concern that the relation between trading volume and liquidity varies across earnings announcement and non-announcement periods, we show in Table A.12 of the Internet Appendix similar correlations between the price impact benchmarks and Amihud measures for earnings announcement period and non-earnings-announcement period.

25 Consensus analyst forecast is the monthly median analyst forecast preceding the earnings announcement. We obtain both consensus forecast and actual earnings from the IBES unadjusted summary file.

Table 9
Turnover premium and earnings announcements, 1972–2012

A. Earnings announcement returns and nonannouncement returns sorted on turnover measures

	Short	2	3	4	Long	L – S	t-stat
Earnings Announcement Return (BHAR [-1, 1])							
Sorted on AT_C	0.02	0.11	0.17	0.27	0.65	0.63***	(7.53)
Sorted on TO	0.01	0.15	0.18	0.31	0.59	0.58***	(6.80)
BHAR for Nonearnings announcement days							
Sorted on AT_C	0.36	0.49	0.38	0.39	0.26	−0.10	(−0.73)
Sorted on TO	0.41	0.51	0.39	0.33	0.22	−0.20	(−1.31)

B. Fama-Macbeth regressions

	Dep. var.: Buy and hold stock return				Dep. var.: Analyst forecast errors	
	Earnings ann. ret.		Non-ann. ret.			
	(1)	(2)	(3)	(4)	(5)	(6)
ln(AT_C)	0.213***		−0.044		−0.089***	
	(7.30)		(−0.74)		(−5.18)	
ln(TO)		−0.240***		0.083		0.106***
		(−7.73)		(1.22)		(6.55)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.014	0.013	0.051	0.050	0.051	0.050
Ave. # obs	579	579	579	579	579	579
# months	492	492	492	492	492	492

Panel A of the table reports the buy and hold returns in the earnings-announcement period and non-earnings announcement-period of portfolios sorted on the turnover measures (*AT_C* and *TO*). At the beginning of each month t from 1972 to 2012, stocks with earnings announcement in the month are sorted into quintile portfolios according to the *AT_C* and *TO* measures of month $t-2$. We define the "long" portfolio as the top (bottom) quintile of *AT_C* (turnover), and the "short" portfolio as the bottom (top) quintile of *AT_C* (turnover). Then for each firm-month, we calculate the buy and hold abnormal return in the three-day window $[-1,1]$ surrounding the earnings announcement, where the buy and hold return is calculated as the buy and hold raw return minus the buy and hold value-weighted CRSP return. BHAR $[-1,1]$ denotes this return. We also calculate the monthly buy and hold return for the days other than the $[-1,1]$ earnings announcement window. In panel A, we first calculate monthly average of BHAR $[-1,1]$ for the quintile portfolios and report time-series average portfolio returns. The differences between the top and bottom quintiles are also reported with associated t-statistics. We also report buy and hold abnormal returns in the nonearnings announcement period instead of BHAR. The t-statistics (in parentheses) are calculated using Newey-West robust standard errors with six lags. In panel B, the left panel presents Fama-Macbeth regressions of BHAR for earnings announcement or nonannouncement period on the turnover measures. The right panel presents Fama-Macbeth regressions of analyst forecast errors on the turnover measures, where analyst forecast error for an announcement is the consensus forecast minus actual earnings, scaled by stock price at the end of the previous quarter. We control for size (*ln(ME)*), book-to-market ratio (B/M), momentum ($Ret[-12,-2]$), reversal ($Ret[-1]$), and an intercept and they are not reported for brevity. t-statistics are calculated using Newey-West robust standard errors with six lags. ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

over-optimism and therefore overpriced than low volume stocks, then we would expect the forecast error to decrease in *AT_C* and increase in turnover. Models (5) and (6) in Table 9, panel B, regress forecast error on *AT_C* and turnover, respectively. The coefficient of *AT_C* is significantly negative, and that on turnover is significantly positive. This result is consistent with Lee and Swaminathan (2000), who find that high volume firms experience lower earnings surprises in the subsequent period. To summarize, the four tests in this section do not support the liquidity premium explanation but are consistent with the mispricing explanation of the volume premium.

5. Pricing of the Amihud Measure as a Systematic Factor

Although our paper focuses on the pricing of the Amihud measure with respect to characteristic liquidity, the Amihud measure has also been used to examine the liquidity commonality (Kamara, Lou and Sadka 2008; Karolyi, Lee, and van Dijk 2012) and the pricing of liquidity as a systematic risk factor (e.g., Acharya and Pedersen 2005). In this section, we extend our analysis to examine whether the trading volume component is also primarily responsible for the pricing of the Amihud measure as a systematic factor.

We create a systematic factor for each Amihud measure: the original Amihud (2002) measure A, the turnover-based Amihud measure AT, and their constant versions (A_C and AT_C), respectively. Following the literature (e.g., Pástor and Stambaugh 2003; Sadka 2006), for each Amihud measure, we obtain the monthly aggregate measure by calculating the equal-weighted average across all stocks, and then estimate time-series regressions using an AR(2) model. We construct the monthly factor as the residual of the AR(2) model multiplied by -1 so that negative values of the factor signify deteriorating market conditions.

We examine the pricing of the systematic factors using the monthly Fama-MacBeth regressions of stock returns on the factor loadings, the level of the Amihud measure (as a characteristic), as well as our standard control variables. The factor loadings are the coefficients of the respective Amihud factors in a firm-level time-series regression using data from month $t-60$ to $t-1$, where the model includes the Fama-French three factors, momentum factor, and the respective Amihud factor. The monthly regressions are estimated from 1967 to 2012, a total of 552 months, and the results are presented in Table 10. We first include our usual controls and then further control for idiosyncratic volatility for robustness test like in our main analysis (Table 3). Consistent with Acharya and Pedersen (2005), we find that the original Amihud measure is priced as a systematic factor. The coefficient for the A factor beta is 0.004 in Model (4) of panel A, which is translated into a monthly premium of 0.17% for one-standard-deviation change in beta (41.86 in our sample). The A_C beta is also priced but the residual A factor beta, the residual of a cross-sectional regression of A factor beta on A_C factor beta, is not priced. The estimated coefficients for the A_C factor beta (0.083) in Model (5) corresponds to a monthly premium of 0.12% for one-standard-deviation change in beta (1.485 in our sample). This magnitude is slightly higher than the premium (0.09%) documented by Acharya and Pedersen (2005).

The results in panel B further provide evidence that the AT_C beta has similar information as the AT beta. When idiosyncratic volatility is not included as a control variable, neither beta has a significant premium, but when idiosyncratic volatility is included, both yield significant premiums and the economic magnitude is similar too. The estimated coefficients in Models (4) and (5) imply that the return premium for one standard deviation increase in AT beta (4.18) and AT_C beta (0.064) are 0.09% and 0.08%, respectively, which are

Table 10
Pricing the Amihud measures as systematic factors

A. Factors using the original Amihud measures

	(1)	(2)	(3)	(4)	(5)	(6)
A factor beta	0.003**			0.004**		
	(1.99)			(2.33)		
A_C factor beta		0.066*	0.068*		0.083**	0.086**
		(1.89)	(1.95)		(2.42)	(2.52)
Residual A factor beta			0.003			0.003
			(0.95)			(0.88)
ln(A)	0.113***	0.114***	0.112***	0.104***	0.105***	0.103***
	(2.78)	(2.78)	(2.76)	(2.69)	(2.70)	(2.66)
Idio. vol.				−13.207**	−13.174**	−13.034**
				(−2.57)	(−2.56)	(−2.55)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.033	0.033	0.035	0.041	0.041	0.043
Ave. # obs	1735	1735	1735	1735	1735	1735

B. Factors using on the turnover-based Amihud measures

	(1)	(2)	(3)	(4)	(5)	(6)
AT factor beta	0.019			0.022**		
	(1.62)			(1.95)		
AT_C factor beta		0.971	0.846		1.213*	1.111*
		(1.41)	(1.20)		(1.83)	(1.65)
Residual AT factor beta			0.018			0.020
			(0.82)			(0.98)
ln(A)	0.115***	0.117***	0.115***	0.106***	0.108***	0.106***
	(2.83)	(2.88)	(2.85)	(2.75)	(2.79)	(2.76)
Idio. vol.				−12.901**	−12.901**	−12.464**
				(−2.53)	(−2.52)	(−2.45)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.033	0.033	0.036	0.041	0.041	0.043
Ave. # obs	1,735	1,735	1,735	1,735	1,735	1,735

This table presents the estimation results of monthly Fama-MacBeth regressions of stock returns on the loadings of factors created using various versions of the Amihud (2002) measure from 1967 to 2012 (552 months). For each Amihud measure, we obtain an aggregate measure by calculating the equal-weighted average of the Amihud measure. We create a factor from the residuals of an AR(2) model on the aggregate measure. We multiply the residual series by −1. The factor beta is the coefficient for the Amihud factor in a firm-level time-series regression using data from month $t-60$ and $t-1$ where the model includes the Fama-French four-factors and the respective Amihud measure factor. Residual A factor beta is the residual of a cross-sectional regression of A factor beta on A_C factor beta. Residual AT factor beta is the residual of a cross-sectional regression of AT factor beta on AT_C factor beta. In addition to ln(A), we also control for *ln(ME)*, B/M, R*et*[−12,-2], and *Ret*[-1]. We estimate a cross-sectional regression in each month and then report the time-series means and t-statistics (in parentheses). t-statistics are calculated using Newey-West robust standard errors with six lags. We also report the time-series average of the number of observations and adjusted R^2 of the cross-sectional regressions. ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

comparable to the premium (0.09%) reported by Acharya and Pedersen (2005). Additionally, when AT_C beta and Residual AT factor beta—the residual of a cross-sectional regression of AT factor beta on AT_C factor beta—are included, the coefficient of the Residual AT factor beta is not significant, whereas the coefficient for the AT_C factor beta remains significant. Overall, the results in Table 10 suggest that the pricing of the Amihud measure as a systematic factor is also primarily driven by the volume component, not its return-to-volume construct.

We also conduct the factor analysis in alternative settings. First, we follow Korajczyk and Sadka (2008) and construct three sets of across-measure common factors for different sets of liquidity measures in 1993–2012, and estimate Fama-Macbeth regressions like in their Table 10.[26] The result, not tabulated for the sake of brevity, shows that including volume is critical to the pricing of across-measure liquidity factor. Specifically, when we replace A with A_C in measure construction, the across-measure systematic factor remains significantly priced. When trading volume is not included, the loading on the across-measure systematic factor does not have a significant premium. Second, Wu (2016) finds that an extreme liquidity risk factor created using the Amihud measure is priced. We follow her empirical framework but construct an alternative measure of extreme risk factor by replacing the Amihud measure with its constant version. In untabulated results, we find similar return premiums for the factor loadings on the two measures of extreme risk.

6. Conclusion

We examine the pricing of the Amihud (2002) measure, one of the most widely used liquidity proxies in the current finance literature. We find that the return premium associated with the Amihud (2002) measure is driven by its association with trading volume, but not its construction of the return-to-volume ratio to capture price impact. A "constant" measure using only the trading volume component exhibits a return predictability matching that of the Amihud (2002) measure, and the return premium associated with the Amihud (2002) measure disappears once the variation of the trading volume component is removed. These findings survive a broad set of robustness tests.

Further analyses show that the high-frequency price impact and the spread benchmarks do not explain the pricing of the trading volume component of the Amihud measure. In fact, the pricing of these liquidity benchmarks exhibits strong January seasonality and disappears outside of January. Additionally, we find evidence that the return premium associated with trading volume is associated with mispricing, but not with the liquidity premium. Finally, we extend the analysis to systematic liquidity factor and the results show that the pricing of the Amihud measure as a systematic factor is also due to its volume component.

Our findings deepen the understanding of the Amihud (2002) measure, a very widely used liquidity measure in the finance literature. On one hand, we confirm that the Amihud (2002) measure does a good job capturing stock liquidity and price impact, as the Amihud measure is highly correlated with the high-frequency price impact benchmark. Therefore, the Amihud measure is useful in measuring the level of stock illiquidity. On the other hand, our findings contradict the general view that the pricing of the Amihud measure captures the compensation for the price impact or liquidity premium. Our findings therefore

[26] We thank Professor Ronnie Sadka for providing the data used in this analysis.

call for caution in the use of the Amihud measure to examine the liquidity premium, control for liquidity in the tests of asset pricing, or construct a liquidity factor.

Our findings also have important general implications for how to measure liquidity and how liquidity affects security prices. Motivated by the rapidly growing literature of stock liquidity, a number of studies have proposed low-frequency liquidity proxies using daily stock market data, and the validity of these measures is usually assessed by whether they are correlated with expected returns. Goyenko, Holden, and Trzcinka (2009) realize this issue and shed light on how well these low-frequency measures measure liquidity by examining their correlations with the corresponding high-frequency liquidity benchmarks.[27] Our findings illustrate the importance of conducting in-depth analysis of the return premium of low-frequency liquidity measure. Additionally, our results show that the price impact and spread benchmarks, the major components of transaction cost, are priced only in January, but not in the full sample period. This puzzling result seems to contradict the theory and calls for further analysis.

Appendix A. Procedures to Clean the Quotes and Trades Data

Following Holden and Jacobsen (2014), we use only NBBO eligible quotes from 9:00 a.m. to 4:00 p.m Eastern Standard Time. For the TAQ data, we do not consider quotes with mode among {4,7,9,11,13,14,19,20,27,28}. For ISSM data, NBBO eligible quotes are those with mode in (‘ ’, ‘A’, ’B’, ‘H’, ‘O’, ‘R’). Quotes meeting one of the filters below are not considered in the NBBO calculation: (1) Bid >offer >0; (2) Bid >0 and offer =0; (3) Offer >0 and bid = 0; (4) Spread >5 and bid >0 and offer >0; (5) Offer if offer <=0 or missing; Offer if size <= 0 or missing; and (6) Bid if bid <=0 or missing; Bid if size <=0 or missing. Note that these invalid quotes are not deleted although they are not considered for the purpose of calculating NBBOs.

We only keep trades in the trading hours from 9:30 a.m. to 4:00 p.m. For ISSM data, trades with special sale condition (‘C’, ‘L’, ‘N’, ‘R’, ‘O’, ‘Z’) are excluded. For the TAQ data, trades with sale condition (’A’ ’C’ ’D’ ’G’ ’L’ ’N’ ’O’ ’R’ ’X’ ’Z’ ’8’ ’9’) are excluded. A trade also needs to have a positive price and size. After cleaning up the quotes and trades data, we then match each trade with the prevailing NBBO. Before 1999, we assume a quote delay of 2 seconds, and zero second between 1999 and 2005. For the year 2006 and afterwards, we adopt the methodology in Holden and Jacobsen (2014). Each trade is then assigned according to the Lee and Ready (1991) trading classification method or the tick test. A trade is classified as buyer (seller) initiated if the trade price is above (below) the prevailing quote midpoint. If the trade price is equal to the midpoint, then we use the tick test.

[27] Corwin and Schultz (2012) is another example of validating low-frequency measures using corresponding high-frequency liquidity benchmarks.

Table A1
Variable definitions

Variable	Definition
A	The original Amihud (2002) measure, constructed as $A_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{\lvert r_{id}\rvert}{Dvol_{id}}$, where r_{id} and $Dvol_{id}$ are daily return and daily dollar trading volume for stock i on day d; D_{it} is the number of days with available ratio in the estimation period t
A_C	The "constant" version of the Amihud measure, constructed as $A\_C_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{1}{Dvol_{id}}$
AT	The turnover-based Amihud measure from Brennan, Huh, and Subrahmanyam (2013), constructed as $AT_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{\lvert r_{id}\rvert}{TO_{id}}$, where TO_{id} is the daily turnover
AT_C	The "constant" version of the turnover-based Amihud measure, $AT\_C_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{1}{TO_{id}}$
Residual A (Res. A) measure	Residuals from the monthly cross-sectional regressions of the A measures on the A_C measures
Residual AT (Res. AT) measure	Residuals from the monthly cross-sectional regressions of the AT measures on the AT_C measures
|Ret|	The return component of the Amihud measure is calculated as the average of daily absolute returns over the estimation period of the Amihud measure
S	The size component of the Amihud measure is calculated as the average of daily market capitalization over the estimation period of the Amihud measure
AN and AP	The two "half" Amihud measures: $AN_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{-\min[r_{id},0]}{Dvol_{id}}$, $AP_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{\max[r_{id},0]}{Dvol_{id}}$
AN_C and AP_C	The "constant" version of the half Amihud measures: $AN\_C_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{-\min[r_{id},0]/r_{id}}{Dvol_{id}}$, $AP\_C_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{\max[r_{id},0]/r_{id}}{Dvol_{id}}$
ATN and ATP	The two "half" turnover-based Amihud measures from Brennan, Huh, and Subrahmanyam (2013): $ATN_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{-\min[r_{id},0]}{TO_{id}}$, $ATP_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{\max[r_{id},0]}{TO_{id}}$
ATN_C and ATP_C	The "constant" version of the half turnover-based Amihud measures: $ATN\_C_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{-\min[r_{id},0]/r_{id}}{TO_{id}}$, $ATP\_C_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{\max[r_{id},0]/r_{id}}{TO_{id}}$
ATNS and ATPB	The two "half and directional" turnover-based Amihud measures constructed using buy volume and sell volume: $ATNS_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{-\min[r_{id},0]}{STO_{id}}$, $ATPB_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{\max[r_{id},0]}{BTO_{id}}$ where BTO and STO are daily buy and sell turnover
ATNS_C and ATPB_C	The "constant" versions of ATNS and ATPB. $ATNS\_C_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{-\min[r_{id},0]/r_{id}}{STO_{id}}$, $ATPB\_C_{it}=\frac{1}{D_{it}}\sum_{d=1}^{Dit}\frac{\max[r_{id},0]/r_{id}}{BTO_{id}}$

(continued)

Table A1
Continued

Variable	Definition
ME	Market capitalization at the end of the previous year
B/M	The book-to-market ratio
Ret[−12,−2]	The cumulative stock return from month $t-12$ to month $t-2$
Ret[−1]	Stock return of month $t-1$
λ	High-frequency price impact measure from Hasbrouck (2009), estimated using the regression model $r_n = \lambda \times SVol_n + u_n$, where for the nth 5-minute period, r_n is the 5-minute stock return calculated as the natural log of the price change over the nth period. (We use quote midpoint instead of trade price to calculate price change.) $SVol_n$ is the signed square root dollar volume of the n^{th} period, and u_n is the error term. We calculate signed square root dollar volume as $SVol_n = \sum_{k=1}^{K_n} sign_k \times \sqrt{dvol_k}$, where $dvol_n$ is the dollar volume of the kth trade in the nth 5-minute period, K_n is the number of trades in the nth period, and $sign_k$ is the sign of the kth trade assigned according to the Lee and Ready (1991) trading classification method or the tick test
Percentage quoted spread (QS)	The difference between the bid and ask quote, divided by the midpoint. The spread is averaged across the estimation period
Percentage effective spread (ES)	The dollar effective spread, $2 \times \lvert P_k - M_k \rvert$, divided by the quotes midpoint, where P_k is the price of the kth trade, and M_k is the prevailing midpoint for the kth trade. The spread is calculated for each trade and then averaged across the estimation period.
Percentage realized spread (RS)	The dollar realized spread, $2 \times Sign_k \times \lvert P_k - M_{k+5} \rvert$, divided by the post-trade quotes midpoint M_{k+5}. M_{k+5} is the prevailing midpoint 5 minutes after the kth trade, and $sign_k$ is the sign of the kth trade assigned according to the Lee and Ready (1991) trading classification method or the tick test. The spread is calculated for each trade and then averaged across the estimation period
Percentage 5-minute price impact (PI)	The dollar effective spread minus the dollar realized spread, scaled by M_{k+5}. The *PI* measure is calculated for each trade and then averaged across the estimation period

References

Acharya, V. V., and L. H. Pedersen. 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.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang. 2006. The cross-section of volatility and expected returns, *Journal of Finance* 61:259–99.

Baker, M., and J. Wurgler. 2006. Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61:1645–80.

Barinov, A. 2014. Turnover: liquidity or uncertainty? *Management Science* 60:2478–95.

Ben-Rephael, A., O. Kadan, and A. Whol. 2015. The diminishing liquidity premium. *Journal of Financial and Quantitative Analysis* 50:197–229.

Blume, L., D. Easley, and M. O'Hara. 1994. Market statistics and technical analysis: The role of volume. *Journal of Finance* 49:153–81.

Brennan, M. J., T. Chordia, and A. Subrahmanyam. 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns, cross-sectional determinants of expected returns, *Journal of Financial Economics* 49:345–73.

Brennan, M. J., S.-W. Huh, and A. Subrahmanyam. 2013. An analysis of the Amihud illiquidity premium. *Review of Asset Pricing Studies* 3:133–76.

Chordia, T., S.-W. Huh, and A. Subrahmanyam. 2009. Theory-based illiquidity and asset pricing. *Review of Financial Studies* 22:3629–68.

Collin-Dufresne, P., and V. Fos. 2015. Do prices reveal the presence of informed trading? *Journal of Finance* 70:1555–82.

Corwin, S. A., and P. Schultz. 2012. A simple way to estimate bid-Ask spreads from daily high and low prices. *Journal of Finance* 67:719–60.

Davis, L. J., E. F. Fama, and K. R. French. 2000. Characteristics, covariances, and average returns: 1929 to 1997. Journal of Finance 55:389–406.

Easley, D., M. Lopez de Prado, and M. O'Hara. 2012. Flow toxicity and liquidity in a high frequency world. *Review of Financial Studies* 25:1457–93.

Easley, D., M. Lopez de Prado, and M. O'Hara. 2016. Discerning Information from Trade Data. *Journal of Financial Economics* 120:269–85.

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

Engelberg, J., and R. D. McLean, and J. Pontiff. 2016. Anomalies and news. Working Paper, Boston College.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81:607–36.

Florackis, C., A. Gregoriou, and A. Kostakis. 2011. Trading frequency and asset pricing on the London Stock Exchange: Evidence from a new price impact ratio. *Journal of Banking and Finance* 35:3335–3350.

Fong, K. Y. L., C. W. Holden, and C. Trzcinka. Forthcoming. What are the best liquidity proxies for global research? *Review of Finance*.

Foster, F. D., and S. Viswanathan. 1993. Variations in trading volume, return volatility, and trading costs: evidence on recent price formation models. *Journal of Finance* 48:187–211.

Gervais, S., R. Kaniel, and D. H. Mingelgrin. 2001. The high-volume return premium. *Journal of Finance* 56:877–919.

Glosten, L. R., and L. E. Harris. 1988. Estimating the components of the bid-ask spread. *Journal of Financial Economics* 21:123–42.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka. 2009. Do liquidity measures measure liquidity? Journal of Financial Economics 92:153–81.

Harris, M., and A. Raviv. 1993. Differences of opinion make a horse race. *Review of Financial Studies* 6:473–506.

Hasbrouck, J. 2009. Trading costs and returns for US equities: Estimating effective costs from daily data. *Journal of Finance* 64:1479–512.

Holden, C. W., and S. Jacobsen. 2014. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. *Journal of Finance* 69:1747–85.

Holden, C. W., S. E. Jacobsen, and A. Subrahmanyam. 2014. The empirical analysis of liquidity. Research Paper, Kelley School of Business, Indiana University.

Jiang, G., C. M. C. Lee, and G. Y. Zhang. 2005. Information uncertainty and expected returns. *Review of Accounting Studies* 10:185–221.

Johnson, T. C. 2008. Volume, liquidity, and liquidity risk. *Journal of Financial Economics* 87:388–417.

Kandel, E., and N. D. Pearson. 1995. Differential interpretations or information and trade in speculative markets. *Journal of Political Economy* 103:831–72.

Kamara, A., X. Lou, and R. Sadka. 2008. The divergence of liquidity commonality in the cross-section of stock. *Journal of Financial Economics* 89:444–66.

Karolyi, G. A., K.-H. Lee, and M. A. van Dijk. 2012. Understanding commonality in liquidity around the world. *Journal of Financial Economics* 105:82–112.

Korajczyk, R., and R. Sadka. 2008. Pricing the commonality across alternative measures of liquidity. *Journal of Financial Economics* 87:45–72.

La Porta, R., J. Lakonishok, A. Shleifer, and R. Vishny. 1997. Good news for value stocks: Further evidence on market efficiency. *Journal of Finance* 52:859–74.

Lee, C. M. C., B. Mucklow, and M. J. Ready. 1993. Spreads, depths, and the impact of earnings information: an intraday analysis. *Review of Financial Studies* 6:345–74.

Lee, C. M. C., and M. J. Ready. 1991. Inferring trade direction from intraday data. *Journal of Finance* 46:733–46.

Lee, C. M. C., and B. Swaminathan. 2000. Price momentum and trading volume. *Journal of Finance* 55:2017–69.

Lou, X., and R. Sadka. 2011. Liquidity level or liquidity risk? Evidence from the financial crisis. *Financial Analysts Journal* 67:51–62.

Miller, E. M. 1977. Risk, uncertainty, and divergence of opinion. *Journal of Finance* 32:1151–68.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–8.

Pástor, L., and R. F. Stambaugh. 2003. Liquidity risk and expected stock returns. *Journal of Political Economy* 111:642–85.

Sadka, R. 2006. Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. *Journal of Financial Economics* 8:309–49.

Stambaugh, R. F., J. Yu, and Y. Yuan. 2012. The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104:288–302.

Wu, Y. 2016. Asset pricing with extreme liquidity risk. Working Paper, Stevens Institute of Technology.