The Night and Day of Amihud's (2002) Liquidity Measure*

Yashar H. Barardehi
Dan Bernhardt
Thomas G. Ruchti
Marc Weidenmier

July 23, 2019

Abstract

Amihud's (2002) stock (il)liquidity measure averages the daily ratio of absolute close-to-close return to dollar volume, including overnight returns, while trading volumes come from regular hours. Our modified measure addresses this mis-match by using open-to-close returns. It better explains cross-sections of returns, doubling estimated liquidity premia. We uncover the mechanism behind this improvement. Using non-synchronous trading near close as an instrument reveals that overnight returns are primarily information-driven and orthogonal to price impacts of trading: including them in liquidity proxies magnifies measurement error, understating liquidity premia. Our modification especially matters for finance/accounting applications that render use of intraday data infeasible/undesirable.

*We are grateful for comments received from Asaf Bernstein, Charles Calomiris, Terry Hendershott, Eric Hughson, Tim Johnson, Andriy Shkilko, Elvira Sojli, and Mitch Warachka. Marc Weidenmier acknowledges Financial support from Chapman University. Any error is our own. Barardehi (barardehi@chapman.edu) is at the Argyros School of Business & Economics, Chapman University. Bernhardt (danber@illinois.edu) is at Department of Economics, University of Illinois and University of Warwick. Ruchti (ruchti@andrew.cmu.edu) is at the Tepper School of Business, Carnegie Mellon University. Weidenmier (weidenmi@chapman.edu) is at the Argyros School of Business & Economics, Chapman University and NBER.

1 Introduction

The stock (il)liquidity measure proposed by Amihud (2002) is the most widely-used such measure in empirical financial economics.[1] Its key advantage stems from its simple construction, which requires only daily return and dollar volume data that are available for many markets and countries over long periods of time. Amihud's measure has consistently produced evidence of priced idiosyncratic liquidity and liquidity risk, and it has been found to be a reasonable proxy of institutional trading costs.[2]

Our contributions are two-fold. First, we identify a simple yet striking improvement to Amihud's measure of liquidity. Our modified measure better explains the cross-section of returns, yet requires almost no incremental data processing. We also document that liquidity premia are larger than previously believed. Second, we uncover the source of this improvement using non-synchronous trading as an instrument. We establish that overnight returns are largely divorced from the price impacts of trading, and are likely information driven. Our correction eliminates noise due to inclusion of these over-night information-driven price movements that would otherwise bias estimated liquidity premia downward.

The widespread use of Amihud's (2002) measure outside of market microstructure underscores the applicability of our proposed correction. Intraday trade and quote data are simply unavailable for many applications, necessitating the use of liquidity measures like Amihud's.[3] In corporate finance and accounting research where use of intraday data is rare, Amhihud's measure is widely-used as an easy-to-construct proxy for liquidity.

[1] Amihud's (2002) article has 1,804 citations by peer-reviewed published articles, including 247 in the top-three finance journals and 50 in the top-three accounting journals (Web of Science, accessed May 6, 2019).

[2] See Chordia et al. (2000), Pástor and Stambaugh (2003), Acharya and Pedersen (2005), Sadka (2006), Asparouhova et al. (2013), Drienko et al. (2017), Harris and Amato (2019), among others. Goyenko et al. (2009) show that Amihud's measure remains priced post-decimalization. Anand et al. (2013) and Barardehi et al. (2019) provide evidence of strong time-series and cross-sectional correlations between Amihud's liquidity measure and actual and estimated institutional trading costs. Studies such as Lipson and Mortal (2004) relate equity liquidity, captured by Amihud's measure, and corporate finance decisions.

[3] Kingsley et al. (2017) find Amihud's (2002) measure to be unsurpassed as a "cost-per-dollar-volume" proxy for global research. Goyenko et al. (2009) show that it is a good proxy for the price impacts of trading.

The standard Amihud (2002) measure, which we abbreviate as $CCAM$ for close-to-close Amihud measure, is constructed using averages of a daily proxy of price impact that divides daily absolute return by the same day's dollar trading volume. The measure captures the price impact of trading, or the amount a given trading volume moves market prices. While the denominator reflects trading volume during trading hours, the numerator reflects close-to-close returns, or absolute return between the close of the previous day and the current day, which includes overnight price movements. Such overnight adjustments reflect after-hours price movements that are often driven by information arrival unrelated to the daily trading volume used in the denominator (Barclay and Hendershott 2003; Santosh 2016).[4]

Our measure, which we abbreviate as $OCAM$ for open-to-close Amihud measure, instead uses the absolute return between opening and closing prices of the trading day, excluding overnight price movements. This modification reflects two observations. First, the vast bulk of trade occurs during trading hours, meaning that trading costs are realized then.[5] This means that a measure better reflects trading costs, or the price impacts of trading volume, if it is calculated using data from trading hours. The same logic underlies the calculation of other measures of stock liquidity such as spreads or estimates of Kyle's λ using trade and quote data exclusively from regular trading hours. Second, trading volumes reported by data vendors such as CRSP almost entirely reflect transactions during trading hours.[6] Figure 1 illustrates the time mismatch between measures of daily return and trading volume that, respectively, enter the numerator and denominator of Amihud's (2002) proxy of daily price impacts. Measuring both the numerator and denominator over regular trading hours keeps these two measurement inputs internally consistent.

[4]Barclay and Hendershott (2003) show that after-hour transactions possess more information content than trades during regular hours. Santosh (2016) finds that 71% of stock value shocks driven by after-hour earnings surprises are reflected in opening prices the following trading day. See also Stole and Whaley 1990 and Cao et al. 2000 for the impact of information revelation at open.

[5]Note that regular hours account for less than 20% of total hours in a week, suggesting that including price movements realized during off-trading hours introduces significant noise.

[6]In recent years, CRSP volume includes the tiny fraction of trades that happen between 4:00pm and 8:00pm on NASDAQ and electronic communication networks (ECNs).

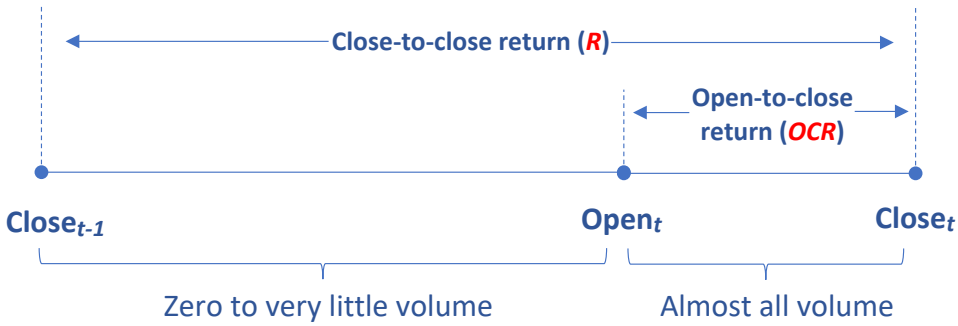

Figure 1: **Illustration of time mismatch between daily measures of trading volume and return.** The figure illustrates the sampling of daily return and trading volume from a typical trading day t. These measures are used to construct Amihud's (2002) daily proxy of price impacts.

We establish that overnight price movements represent a major source of the observed cross-sectional and time-series variations in the close-to-close Amihud measure ($CCAM$). We show that these overnight returns are primarily information-driven, and have little to no relevance for liquidity measurement. Including them in the illiquidity measure introduces measurement error that could bias estimates of liquidity premia toward zero. $OCAM$ excludes such price movements, producing significantly higher explanatory power for the cross-section of returns, and revealing larger liquidity premia.

Figure 2 reveals the impact of including overnight price movements on the measurement of liquidity. Were there negligible differences between the two Amihud measures, then the ratio of $OCAM$ to $CCAM$ would be close to one. Figure 2 plots the three cross-stock quartiles of this ratio over time. Several key findings emerge: (i) the ratio is far less than one for the bottom quartile of stocks in the first thirty years of the sample (ranging from below 0.5 to about 0.7 before rising to 0.8 and above by the mid-1990s), a pattern that cannot be explained by cross-stock variation in after-hours trading;[7] (ii) cross-stock variation appears to decline over time; (iii) while the ratio displays a generally positive trend, there are several

[7] After-hours trading did not exist until 1991, and it was so thin that it motivated research to explain this anomaly (e.g., Belcourt (1996)). Moreover, years before after-hours trading on ECNs was introduced, disparities between the two measures had already fallen sharply (close to current levels).

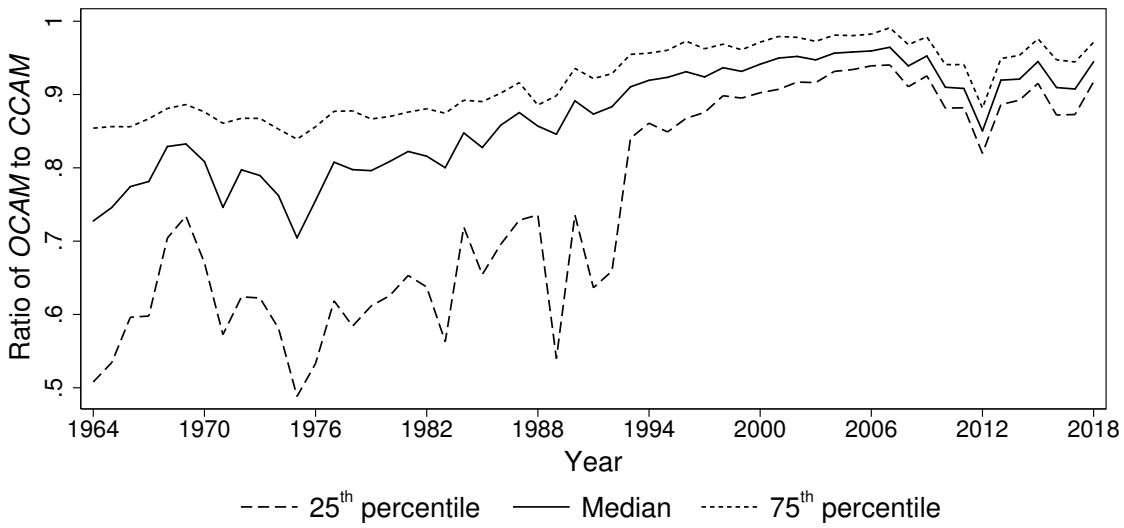

Figure 2: **Evolution of the ratio of $OCAM$ to $CCAM$.** The figure plots the temporal changes in the cross-stock distribution of the $OCAM$-to-$CCAM$ for NYSE-listed common shares in the 1964-2017 period. For each stock i in year y, the ratio $ROC_{iy} = OCAM_{i,y-1}/CCAM_{i,y-1}$ is calculated. The year-specific quartiles of this ratio are plotted against time. See Section 2.2 for variable construction.

episodes of sharp systematic decline such as that during the 2008 financial crises, highlighting the economic relevance of $OCAM$ even in modern financial markets; (iv) the ratio is less than one for almost 95% of stock-years; and (v) reflecting the temporal stability of the ratio, the two measures display a strong time-series co-movement, indicating that $OCAM$ likely picks up temporal variation in price impacts similarly to $CCAM$.

Our modification leads to improvements in the ability of Amihud's measure to explain the cross-section of returns. We replicate and extend the cross-sectional analysis of Amihud (2002) for NYSE-listed firms in the 1964–2017 period, using a novel database that reports historical open prices. Liquidity premia based on $OCAM$ are roughly double those obtained based on the traditional $CCAM$—in Appendixes B and C, we verify the robustness of this finding to corrections for upward biases identified by Asparouhova et al. (2010, 2013).[8] We

[8]Also, because we standardize both measures for these tests, it is not the fact that $OCAM$ is smaller in magnitude that is driving our findings.

document similar evidence when we use the two measures to estimate liquidity premia on risk-adjusted returns from a Fama-French four-factor model in panel regressions. These estimates control for both stock and time fixed effects, revealing that neither unknown fixed stock characteristics nor temporal variation in sample composition drive our findings. Temporal patterns in liquidity premia are consistent with findings of Asparouhova et al. (2010), Ben-Rephael et al. (2013), and Harris and Amato (2018) that liquidity premia decline over time.

Our findings are robust to several considerations. First, we establish that temporal changes in the composition of listed stocks do not explain the disparity between $CCAM$ and $OCAM$. Second, estimates are unaffected if we eliminate any impacts of open auctions by excluding intervals at open. Third, qualitatively identical outcomes obtain when we exclusively rely on CRSP data to construct $OCAM$. Indeed, using the CRSP-based 1993–2017 sample (CRSP first reports open prices in 1992), we find that $OCAM$ produces statistically significant liquidity premia of 3.7–4.8 bps per month for NYSE- and AMEX-listed stocks, but $CCAM$ is not associated with significant liquidity premia. Fourth, analysis of AMEX- and NASDAQ-listed stocks reveals that estimated liquidity premia using $OCAM$ are again roughly double those using $CCAM$. In fact, liquidity premia for (the generally smaller) AMEX- and NASDAQ-listed stocks are two to five times higher than those for NYSE-listed firms, and they remain high in recent years.

To evaluate the incremental information content of each measure, we decompose each version of Amihud's (2002) measure into linearly orthogonal components with respect to the alternative. The residuals from regressing $OCAM$ on $CCAM$ significantly explain the cross-section of expected (risk-adjusted) returns. However, the converse is not true. We show in Appendix E that $OCAM$'s superior performance in capturing liquidity, relative to $CCAM$, also manifests itself in the higher correlations of $OCAM$ with other standard measures of stock liquidity.

Our findings also inform the debate on the pricing of the Amihud (2002) liquidity mea-

sure. Lou and Shu (2017) argue that the pricing of Amihud's (2002) measure is driven by variation in the trading (dollar) volumes entering its denominator as opposed to variation in absolute close-to-close returns that enter the numerator.[9] Our analysis provides direct evidence that the absolute return component (the numerator) drives pricing. Our modification alters the absolute return component by removing overnight price movements, leaving the dollar-volumes in the denominator unchanged. This single change sharply alters pricing of Amihud's measure in the cross-section.

We complete our analysis by uncovering what drives the cross-sectional and temporal variation in the $OCAM/CCAM$ ratio. In particular, we explain its rise and convergence post decimalization, establishing that the improvements $OCAM$ makes over $CCAM$ reflect (overnight) information-driven price movements. To do this, we exploit non-synchronous trading near close, measured by the time distance between the last transaction of the day and 4:00pm EST.[10] If a stock experiences a longer period of non-trading before close and information flows per unit time are similar across stocks, then it should see more accumulated information after the final trade, but before close. That accumulated information is impounded into price at open on the next trading day, leading to greater overnight price movement.

The cross-sectional variation in the extent of non-trading across stocks explains a remarkable share of the cross-sectional variation in the ratio of $OCAM$ to $CCAM$ post-1993. Indeed, after accounting for the extent of non-trading, the cross-sectional distribution of the $OCAM/CCAM$ residual becomes stable over time—the temporal variation in the distribution of the ratio is almost entirely explained by that in the distribution of non-trading. After controlling for stock size and turnover, along with stock and time fixed effects, each

[9] Amihud and Noh (2018) argue that the log-linearization used by Lou and Shu (2017) to decompose Amihud's (2002) measure into the sum of the log of absolute return and the log of one over dollar volume is incorrect. Lou and Shu (2017) omit the correlation term between daily absolute return and the inverse of dollar volume, which Amihud and Noh (2018) show is priced.

[10] The literature has previously identified the importance of non-synchronous trading for large autocorrelations in the returns of stock indexes. See Atchison et al. (1987).

additional ten minutes of non-trading prior to close is associated with a 5-percentage-point decline in the $OCAM/CCAM$ ratio. Thus, extensive differences in non-trading near close imply large disparities between the two liquidity measures. This analysis ties the extent of greater measurement error in $CCAM$ vis à vis $OCAM$ to information-driven price movements that are unrelated to price impacts of trading.

Our paper contributes to a literature that distinguishes price movements during trading hours from those when markets are closed (e.g., Cliff et al. 2008; Barclay and Hendershott 2008; Hendershott et al. 2018; see also French and Roll 1986). Our modification highlights the importance of time-matched price movement and trading volume inputs for Amihud's liquidity measure. Using this modification reveals that liquidity premia are larger than previously believed. We document a strong relationship between improvements in the quality of Amihud's measure obtained by removing overnight price movements, and the extent of non-trading before close. This result highlights the information-driven nature of overnight price movements, indicating a broader reach of our analysis.

Post-decimalization, empirical analyses of expected returns in U.S. markets suggest that they have become more liquid; and high frequency liquidity measures are now available. Nonetheless, Amihud's (2002) liquidity measure remains vital for measuring liquidity in international markets.[11] Logic suggests that given how our correction matters for the pricing of liquidity in the U.S.—the world's leading financial market—it would matter more in less liquid international or emerging markets that are fully or partially closed during U.S. trading hours. Indeed, in untabulated results, we find that the median $OCAM/CCAM$ ratio in

[11] Amihud's measure is heavily used in studies of markets outside North America. Amihud et al. (2015) document positive liquidity premia across 45 countries in the 1990−2011 period. Lee (2011) documents priced liquidity risk in international markets. Hung et al. (2014) find weaker post-earnings announcement drifts in international markets that are more liquid according to Amihud's measure. Boehmer et al. (2015) show that increased algorithmic trading across 42 international markets impacts Amihud's (2002) measure. Chen et al. (2017) document eviiidence in developing markets that the initial enforcement of insider trading laws is more effective for firms whose stocks witness improvements in their Amihud (2002) liquidity measures. Lang et al. (2015) find that the Amihud (2002) measure is highly correlated with textual lengths of financial reports for firms in the 42 countries studied.

the Brazilian stock market is roughly half its NYSE counterpart in the 2008–2018 period, indicating the heightened economic importance of our correction for such markets.

2 Data and variables

2.1 Amihud illiquidity and modified Amihud

The traditional Amihud (2002) stock liquidity measure, dubbed $CCAM$ (for close-to-close Amihud), calculates the average of the daily absolute return per dollar traded over a given time period spanning D consecutive trading days, where daily returns reflect close-to-close returns, incorporating overnight price adjustments and dividend distributions. As in Amihud (2002), $CCAM_{iy}$ uses D_{iy} daily observations of stock i in year y,

$$CCAM_{iy} = \frac{1}{D_{iy}} \sum_{d=1}^{D_{iy}} \frac{|R_{idy}|}{DVOL_{idy}}, \tag{1}$$

where R_{idy} and $DVOL_{idy}$, respectively, are stock i's return and dollar trading volume on day d in year y; and D_{iy} is the number of days for which trading volume for stock i in year y is non-zero. In our rolling regression analyses, we use an alternative construction that updates measures monthly. Thus, instead of using annual averages that we indexed by y in (1), we average the daily absolute return per dollar traded over the 12 months ending in month t:

$$CCAM_{it} = \frac{1}{D_{it}^{t-12}} \sum_{d=1}^{D_{it}^{t-12}} \frac{|R_{idt}^{t-12}|}{DVOL_{idt}^{t-12}}. \tag{2}$$

As such, the measure in month t uses stock i's D_{it}^{t-12} daily observations, for days with non-zero trading volume, in the previous 12 months rather than from the previous calendar year.

Our open-to-close version, dubbed $OCAM$, instead uses the open-to-close absolute return to construct daily absolute returns per dollar traded. Stock i's open-to-close return on day d is

$$OCR_{id} = \frac{P_{id}^{c}}{P_{id}^{o}} - 1, \tag{3}$$

where P^o_{id} and P^c_{id} are the open and close prices, respectively. As such, the analogues to the traditional Amihud's measures defined in equations (1) and (2) are

$$OCAM_{iy} = \frac{1}{D_{iy}} \sum_{d=1}^{D_{iy}} \frac{|OCR_{idy}|}{DVOL_{idy}} \tag{4}$$

and

$$OCAM_{it} = \frac{1}{D^{t-12}_{it}} \sum_{d=1}^{D^{t-12}_{it}} \frac{|OCR^{t-12}_{idt}|}{DVOL^{t-12}_{idt}}. \tag{5}$$

2.2 Data and variable definitions

Our main sample runs from January 1, 1963 to December 31, 2017, and contains trade and price information, focusing on all NYSE-listed stocks as in Amihud (2002). In robustness analyses, we extend the sample to include AMEX- and NASDAQ-listed stocks. We obtain daily closing prices, trading volumes, and dividend distributions from Daily CRSP. We match these daily observations with open prices obtained from Global Financial Data (GFD).[12] For stock-days with open price observations in GFD, we match daily observations across CRSP and GFD using security identifiers PERMNO, NCUSIP, and CIK.[13] We obtain monthly returns, prices, dividend distributions, and number of shares outstanding from Monthly CRSP.[14] We match these monthly data with 1-month T-bill rates, and Fama-French four-factor returns from WRDS. We exclude a stock-year set of observations if that stock's daily closing price is below $1 on any day in the preceding calendar year.[15] We follow Amihud (2002) by excluding stock-month observations that are among the 1% least liquid in each month, according to Amihud's measure.

[12] As a robustness check of GFD opening prices, we estimate our asset pricing model for 1993-2017 using opening price data from CRSP, establishing that our findings are unaffected by the data source.

[13] To control for potential data errors in CRSP or GFD, we use similarity in closing prices reported by CRSP and GFD, dropping a matched stock-day observation if its CRSP closing price deviates from that in GFD by more than 0.1%. For example, a stock day with a CRSP closing price of $20.03 and a GFD closing price of $20 is dropped.

[14] We replace a stock's monthly return with its de-listing return `dlret` when a stock is de-listed.

[15] In analyses based on replicating Amihud's (2002) findings, we replace the "penny-stock" filter with one that excludes stocks with end-of-previous-year's closing prices below $5, as in Amihud (2002).

We construct cross-sections of stock characteristics and merge them with cross-sections of monthly returns in two ways. First, following Amihud (2002), we calculate market betas of size portfolios (deciles of market capitalizations at the end of the previous year), β_{py}^{mkt}, using daily stock and equally-weighted market returns every year—we use β_{py}^{mkt} for β_{iy}^{mkt} if stock i is in portfolio p in year y. We then compile the following stock-specific measures at annual frequencies: Dividend yield, DYD_{iy}, is defined as the ratio of total dividend distributions in a year divided by the closing price at the end of the year. Annual measures of momentum are the returns over the last 100 days of the year, $R100_{iy}$, and the realized return over the earlier remaining days of the year, $R100YR_{iy}$. Annual return volatility is captured by the annual standard deviation of daily returns per year, $SDRET_{iy}$. Market capitalization, M_{iy}, is the product of shares outstanding and the closing price at the end of the year. We match these annual measures with each of the monthly return observations of the relevant stock over the following year, to construct an unbalanced monthly panel.

Our second approach addresses the possibility that using the same annual measures of stock characteristics to explain returns in each of the 12 monthly cross-sections in the following year adds noise to later month observations in that year—the previous year's measure grows less germane. This leads us also to use a rolling regression approach, constructing stock characteristics at monthly frequencies, and then matching them with monthly returns in the next month.[16] Fama-French four-factor betas ($\beta_{i,t-1}^{mkt}$, $\beta_{i,t-1}^{hml}$, $\beta_{i,t-1}^{smb}$, and $\beta_{i,t-1}^{umd}$) at the end of month $t-1$ are estimated at the stock level using weekly data from the preceding two years (months $t-24$ to $t-1$), requiring at least one year of data. We use four-factor betas, 1-month T-bill rate, and monthly Fama-French factor portfolio returns to construct risk-adjusted returns. Dividend yield, $DYD_{i,t-1}$, divides total dividend distributions between months $t-12$ and $t-1$ by the closing price at the end of month $t-1$. Momentum measures $RET_{i,t-1}^{t-4}$ and $RET_{i,t-5}^{t-12}$, respectively, capture compound returns over the preceding three months and the

[16] See, e.g., Lu and Shu (2017) and Barardehi et al. (2019).

nine months before that. Return volatility $SDRET_{i,t-1}$ is given by the standard deviation of daily stock returns over the preceding 12 months. Market capitalization, $M_{i,t-12}$, is the product of shares outstanding and the closing price at the end of the month, a year earlier.

We use monthly TAQ data to collect time stamps of the last transaction per trading day, during regular trading hours (9:30am−4:00pm EST), for all NYSE-listed stocks in the 1993−2013 period.[17] We use the temporal distance between these time stamps and 4:00pm EST, in hours, to construct a measure of the extent of non-trading by stock-year. For a given year, we match these observations with our main sample, described above, using NCUSIP from CRSP and CUSIP from TAQ. For observations without such links, we match SYMBOL from TAQ with TSYMBOL from CRSP.

3 Liquidity, overnight returns, and stock attributes

Figure 2 on page 3, which plots the temporal evolution in the distribution of the ratio of open-to-close and close-to-close Amihud (2002), $ROC_{iy} = \dfrac{OCAM_{iy}}{CCAM_{iy}}$, by quartile, reveals that overnight price movements play a major role in driving both cross-sectional and temporal variation in the traditional Amihud (2002) measure, $CCAM$. Including overnight price movements also inflates the measure for most stocks. Importantly, as the evolution of the three quartiles of this ratio indicate, the contamination driven by overnight price movements is time-varying and declining, but not disappearing. Table 1 presents medians of several stock characteristics across six ROC categories, after sorting the sample on ROC on a year-by-year basis. Save for market capitalization, stocks in different such categories do not appear to possess materially different characteristics, indicating that the cross-stock variability in the mismeasurement of liquidity is unlikely to be entirely driven by stock characteristics.

One might wonder whether the patterns documented in Figure 2 could reflect temporal variation in the composition of common stocks over our long sample period. One could posit

[17] The TAQ database is available for 1993 and after.

Table 1: **Stock characteristics by levels of ROC.** This table presents medians of stock characteristics by different categories of ROC, i.e., the ratio of $OCAM$-to-$CCAM$. Each year, stocks are sorted into 6 ROC percentile categories: less than 5%, 5%−25%, 25%−median, median−75%, 75%−95%, and greater than 95%. Medians of stock characteristics, based on observations from the previous year, are calculated by these categories. β^{mkt} is market beta, M is market capitalization in millions of dollars, DYD is dividend yield (%), $SDRET$ is daily return volatility (%), and PRC is the end-of-year closing price.

	Percentile of ROC					
	<5%	5%−25%	25%−50%	50%−75%	75%−95%	>95%
ROC	0.26	0.64	0.85	0.90	0.94	1.02
β^{mkt}	0.99	1.01	1.03	1.03	1.02	0.99
M	315.6	190.5	399.6	653.6	685.1	424.3
DYD	2.30	1.98	1.92	1.99	2.04	1.84
$SDRET$	1.79	2.00	2.12	2.11	2.13	2.22
PRC	26.38	23.00	25.40	27.13	24.75	17.25

that the disparity between $OCAM$ and $CCAM$ might capture some unknown stock characteristics, and that the presence of stocks with small $ROC_{iy} = \dfrac{OCAM_{iy}}{CCAM_{iy}}$ in the cross-section has varied over time. In fact, the number of publicly listed firms varies significantly over the past few decades (Kahle and Stulz 2017).[18] Our panel regression analyses addresses this directly by including stock and time fixed effects. To further preclude the possibility that results are driven by changes in sample composition, we show that the temporal variation in the cross-section of ROC is robust to sample composition. Figure 3 focuses on the sample featuring the 800 stocks with the largest market-capitalizations at the end of the previous year. The figure reveals qualitatively identical patterns in the year-specific cross-stock quartiles of ROC_{iy} to those in Figure 2; there is only a moderate upward shift in early decades in the bottom quartile and median, with minimal shifts for the top quartile (less than 2 percentage points). This finding is consistent with the result in Table 1 that ROC is largely unrelated to several key stock characteristics. Moreover, we rule out the possibility that

[18]The number of stocks per year in our final sample of common NYSE-listed stocks has a minimum of 859 in 1991 and a maximum of 1,370 in 1999. The mean and median number of such stocks are 1,067.7 and 1,070, respectively.

price movements and trading volumes associated with open auctions alter our findings. In Appendix A, we exclude the returns and volumes realized in the first several minutes of the trading day to show that open auctions do not drive our results.

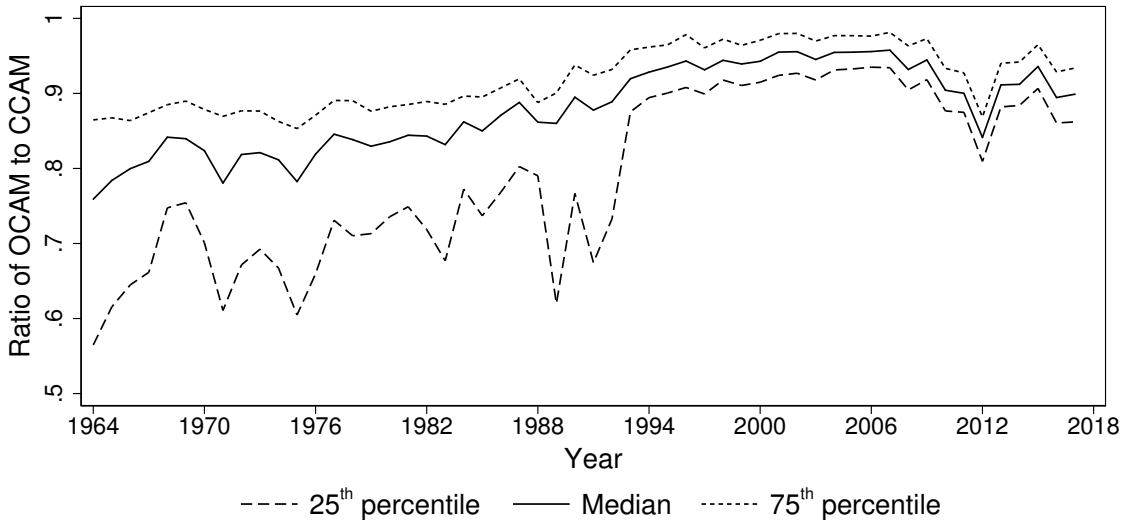

Figure 3: **Evolution of the ratio of $OCAM$ to $CCAM$ (800 largest firms).** The figure plots the temporal changes in the cross-stock distribution of the $OCAM$-to-$CCAM$ among the 800 NYSE-listed common shares with largest market-capitalizations in the 1964-2017 period. Each year stocks are sorted by market-capitalization to obtain the 800 firms featuring the largest market values. For each stock i in year y, the ratio $ROC_{iy} = OCAM_{i,y-1}/CCAM_{i,y-1}$ is calculated. The year-specific quartiles of this ratio are plotted against time.

Because our focus is on the measurement of liquidity, the patterns in Figure 2 are especially relevant for smaller and more thinly-traded stocks, which are generally perceived as relatively less liquid. We document evidence of this by examining how the ratio of $OCAM_{iy}$ to $CCAM_{iy}$, ROC_{iy}, varies with measures of size and turnover. We measure stock i's size in year y by the natural log of its market capitalization at the end of year $y-1$. Our turnover measure for stock i is the natural log of its average daily turnover[19] in year $y-1$, i.e., $\ln(TR_{i,y-1})$. We fit a panel of annual ROC measures against these stock characteristics for the time pe-

[19]Daily turnover is defined as the ratio of the number of shares traded daily to the corresponding number of shares outstanding.

riod 1964−2017, clustering standard errors at both stock and year levels to account for the possibility of inflated t-statistics driven by auto-correlated error terms (see Petersen 2009). We model the cross-sectional variation in ROC_{iy} using the following specification.

$$ROC_{iy} = \alpha_0 + \alpha_1 \ln(M_{i,y-1}) + \alpha_2 \ln(TR_{i,y-1}) + \text{fixed effects} + \epsilon_{iy}, \tag{6}$$

where we include both stock and year fixed effects. Because $OCAM$ is closer to $CCAM$ when ROC_{iy} is larger, a positive coefficient α_1 or α_2 means that the two measures differ less when market capitalization or turnover, respectively, are larger.

Table 2: **Association between $OCAM$-to-$CCAM$ ratio and stock characteristics.** This table presents panel regression estimates when the ratio of open-to-close and close-to-close Amihud (2002) measures, ROC_{iy}, is regressed on the natural logs of market-capitalization, $\ln(M_{i,y-1})$, and mean daily turnover, $\ln(TR_{i,y-1})$, as in equation (6), in the 1964−2017 sample. Specifications differ in the set of fixed effects introduced. Numbers in parentheses reflect standard errors of estimates, clustered at both stock and year levels. Symbols *, **, and *** reflect statistical significance at 10%, 5%, and 1% type one error, respectively.

Dep. Var. = ROC	1964−2017				1964−1980	1981−2000	2001−2017
$\ln(M)$	0.028***	0.038***	0.024***	0.023***	0.036***	0.036***	−0.007***
	(0.003)	(0.004)	(0.004)	(0.005)	(0.016)	(0.006)	(0.002)
$\ln(TR)$	0.064***	0.052***	0.060***	0.058***	0.110***	0.071***	0.008*
	(0.004)	(0.005)	(0.006)	(0.006)	(0.010)	(0.006)	(0.004)
Stock fixed effects	No	Yes	No	Yes	Yes	Yes	Yes
Year fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
Observations	58,120				18,051	21,618	18,451
R^2	0.11	0.32	0.14	0.34	0.27	0.45	0.57

Table 2 shows large positive and significant relationships between ROC and measures of stock size and turnover, indicating that potential "contamination" by overnight price movements is more pronounced for smaller and more-thinly traded stocks. Variation in these two stock characteristics explain over 11% of the variation in ROC, and the positive associations between the ratio and these characteristics remain even in the presence of both stock and year fixed effects.

The fact that the impact of overnight price movements on measures of liquidity is greater for less liquid stocks suggests that our modified measure of liquidity may be priced differently in the cross-section of stock returns. To reinforce this conjecture, we establish that the differences between $CCAM$ and $OCAM$ reflect substantially different cross-sectional rankings of stocks based on the two measures. We find that the differences do not represent a simple scaling effect: rank correlation statistics between $CCAM$ and $OCAM$ are well below one, especially for less liquid stocks and in earlier years of the sample. To quantify this, we sort stocks each year into top 30%, middle 40%, and bottom 30% liquidity according to $CCAM$. We then calculate Kendall's τ statistics every year within each liquidity group, and calculate the average statistic across different years in the entire sample period or in a sub-period.

The two measures of liquidity order stocks in the cross-section very differently. Table 3 shows that rank correlations over the entire sample period are far less than one. These correlations are much smaller for less liquid stocks, going from 76% for the least liquid stocks to 88.9% for the most liquid ones. Consistent with the patterns presented in Figure 2, rank correlation statistics across all liquidity groups substantially rise over time.[20] These findings underscore that $CCAM$ and $OCAM$ measure cross-sectional differences in stock liquidity differently, and hint at potentially different pricing of the two measures in the cross-section.

Table 3: **Rank correlation statistics between $OCAM$ and $CCAM$.** This table presents Kendall's τ statistics across $CCAM$ and $OCAM$ over time periods, 1964−2017, 1946−1980, 1981−2000, and 2001−2017. Each year, stocks are sorted into top 30%, middle 40%, and bottom 30% liquidity according to $CCAM$. Kendall's τ statistic is calculated every year within each liquidity group, and then averaged across different years in the entire sample period or in a sub-period.

Liquidity group	1964–2017	1964–1980	1981–2000	2001–2017
Top 30% liquid	88.9%	87.6%	84.0%	96.0%
Middle 40% liquid	84.9%	72.9%	86.0%	95.7%
Bottom 30% liquid	76.0%	57.7%	74.6%	95.9%

[20] Unreported analyses verify qualitatively similar patterns based on Pearson and Spearman correlations.

4 Modified Amihud measure and liquidity premia

We begin our asset pricing tests by confirming basic results in Amihud (2002) as a benchmark. In particular, using both the classic close-to-close price impact measure from Amihud (2002), $CCAM$, and our proposed open-to-close modified Amihud measure, $OCAM$, we replicate the sixth, eighth, and ninth columns of Table 1 (p. 41) in Amihud (2002) as closely as possible.[21] We then contrast findings based on $CCAM$ with those obtained using $OCAM$ to highlight the value of our modification. We model the cross-section of returns in month t of year y, R_{ity} as in Amihud (2002), by estimating

$$R_{ity} = k^0_{ty} + \sum_{j=1}^{J} k^j_{ty} X^j_{i,y-1} + \epsilon_{ity} \tag{7}$$

using the Fama-MacBeth approach. $X^j_{i,y-1}$ is stock i's j^{th} characteristic, measured using data from year $y-1$; k^j_{ty} is the jth characteristic's loading; and ϵ_{ity} is an error term. To generate estimates that can be compared to those in Amihud (2002), we similarly fit the model for three time periods: 1964−1997, 1964−1980, and 1981−1997.[22] Also following Amihud (2002), we divide each $CCAM_{iy}$ and $OCAM_{iy}$ observation by its respective sample mean across stocks in year y. This centers each liquidity measure to have a mean of one. Thus, the coefficients on liquidity measures reflect liquidity premia: the additional return investors require for holding the stock with average liquidity, compared to the idealized, fully-liquid stock. This centering also ensures that any differences that we find are not mechanically driven by the fact that $OCAM$ is on average smaller than $CCAM$.

Table 4 shows that when we use $CCAM$ to explain cross-sections of stock returns, our qualitative findings perfectly align with those in Amihud (2002). In particular, we find positive and significant coefficients on stock (il)liquidity and measures of momentum, but

[21] We cannot perfectly replicate the results because GFD does not report open prices for virtually all stocks covered by CRSP, and also because for some stocks we cannot identify a valid CRSP-GFD link.

[22] For these estimates, we implement the same filter as in Amihud (2002) for penny stocks. That is, we exclude a stock from year y cross-sections if its closing price at the end of year $y-1$ falls below \$5.

Table 4: **Replication of Table 1 from Amihud (2002) with and without correcting for overnight price movements.** This table presents Fama-MacBeth estimates of equation (7) for all NYSE-listed stocks for time periods 1964−1997, 1964−1980, and 1981−1997. The dependent variable is the monthly stock return in percentage points. $CCAM$ is the traditional Amihud's liquidity measure, and $OCAM$ is Amihud's measure after removing overnight price movements. We divide each $CCAM_{iy}$ and $OCAM_{iy}$ observation by its respective sample mean across stocks in year y, thereby centering each measure to have a mean of one, making coefficients across the two measures comparable. β^{mkt} is market beta estimated across ten size portfolios using daily observations from the last calendar year. $R100$ is the compound return on a stock in the last 100 days of the previous calendar year, and $R100YR$ is the compound return over the earlier remaining trading days in the last calendar year. $\ln(M)$ is the natural log of market capitalization at the end of the previous calendar year. $SDRET$ is the standard deviation of daily returns over the previous calendar year. DYD is the ratio of total cash dividend distribution over the previous calendar year to the closing price at the end of that year, or dividend yield. Newey-West standard errors using two lags are reported in parentheses. Symbols *, **, and *** reflect statistical significance at 10%, 5%, and 1% type one error, respectively.

	1964−1997		1964−1980		1981−1997	
β^{mkt}	−0.280	−0.171	−0.522	−0.506	−0.038	0.164
	(0.308)	(0.306)	(0.442)	(0.437)	(0.429)	(0.424)
$CCAM$	0.056**		0.066*		0.046**	
	(0.022)		(0.039)		(0.021)	
$OCAM$		0.120***		0.157***		0.084***
		(0.032)		(0.055)		(0.031)
$R100$	0.794***	0.774***	0.722**	0.697**	0.867***	0.851***
	(0.227)	(0.227)	(0.332)	(0.332)	(0.310)	(0.309)
$R100YR$	0.389***	0.386***	0.452**	0.457***	0.325**	0.316**
	(0.109)	(0.110)	(0.174)	(0.175)	(0.131)	(0.131)
$\ln(M)$	−0.063	−0.042	−0.149**	−0.117**	0.023	0.033
	(0.040)	(0.038)	(0.062)	(0.057)	(0.050)	(0.051)
$SDRET$	−1.061**	−1.152**	−0.493	−0.612	−1.628**	−1.692**
	(0.495)	(0.509)	(0.709)	(0.742)	(0.684)	(0.690)
DYD	−0.993	−0.975	−1.905	−1.862	−0.082	−0.087
	(1.161)	(1.148)	(2.289)	(2.262)	(0.383)	(0.381)

negative and significant coefficients on stock size and return volatility measures. Substituting $OCAM$ for $CCAM$ leaves the coefficients on other stock characteristics essentially unchanged, but leads to liquidity premia that are more that double those obtained using $CCAM$.[23] The same qualitative findings obtain when we expand the sample period to 1964-

[23]Note that because we normalize $CCAM$ and $OCAM$ to have the same mean of 1, the fact that we find

Table 5: **Fama-MacBeth estimates of monthly returns on stock characteristics, all NYSE-listed stocks, 1964−2017.** This table presents Fama-MacBeth estimates of Equation (7) for all NYSE-listed stocks in time periods 1964−2017, 1964−1980, 1981−2000, and 2011−2017. The dependent variable is the monthly stock return in percentage points. $CCAM$ is the traditional Amihud's liquidity measure, and $OCAM$ is Amihud's measure after removing overnight price movements. We divide each $CCAM_{iy}$ and $OCAM_{iy}$ observation by its respective sample mean across stocks in year y, thereby centering each measure to have a mean of one. β^{mkt} is market beta estimated across ten size portfolios using daily observations from the last calendar year. $R100$ is the compound return on a stock in the last 100 days of the previous calendar year, and $R100YR$ is the compound return over the earlier remaining trading days in the last calendar year. $\ln(M)$ is the natural log of market capitalization at the end of the previous calendar year. $SDRET$ is the standard deviation of daily returns over the previous calendar year. DYD is the ratio of total cash dividend distribution over the previous calendar year to the closing price at the end of that year, or dividend yield. Newey-West standard errors using two lags are reported parentheses. Symbols *, **, and *** reflect statistical significance at 10%, 5%, and 1% type one error, respectively.

	1964−2017		1964−1980		1981−2000		2001−2017	
β^{mkt}	0.047	0.141	−0.522	−0.506	0.272	0.436	0.342	0.434
	(0.229)	(0.228)	(0.442)	(0.437)	(0.423)	(0.416)	(0.306)	(0.312)
$CCAM$	0.043^{***}		0.066^{*}		0.042^{**}		0.021^{**}	
	(0.014)		(0.039)		(0.019)		(0.010)	
$OCAM$		0.090^{***}		0.157^{***}		0.076^{***}		0.040^{***}
		(0.021)		(0.055)		(0.029)		(0.014)
$R100$	0.586^{***}	0.569^{***}	0.722^{**}	0.697^{**}	0.671^{**}	0.654^{**}	0.356	0.349
	(0.165)	(0.165)	(0.332)	(0.332)	(0.290)	(0.289)	(0.215)	(0.214)
$R100YR$	0.270^{***}	0.268^{***}	0.452^{**}	0.457^{***}	0.283^{**}	0.274^{**}	0.079	0.078
	(0.086)	(0.086)	(0.174)	(0.175)	(0.140)	(0.141)	(0.131)	(0.130)
$\ln(M)$	-0.064^{**}	-0.049^{*}	-0.149^{**}	-0.117^{**}	0.015	0.026	-0.072^{**}	-0.070^{**}
	(0.028)	(0.027)	(0.062)	(0.057)	(0.048)	(0.048)	(0.031)	(0.031)
$SDRET$	-0.761^{*}	-0.824^{**}	−0.493	−0.612	-1.540^{**}	-1.596^{**}	−0.130	−0.148
	(0.395)	(0.403)	(0.709)	(0.742)	(0.647)	(0.653)	(0.695)	(0.697)
DYD	−1.061	−1.045	−1.905	−1.862	0.008	0.009	-1.462^{**}	-1.455^{**}
	(0.748)	(0.739)	(2.289)	(2.262)	(0.332)	(0.330)	(0.565)	(0.566)

2017 in Table 5. We find the same doubling in the estimated liquidity premium, even in recent years when differences in the magnitudes of $CCAM$ and $OCAM$ are small.

The late 1990s and the new millennium featured a massive increase in exchange-traded funds, whose shares are traded on U.S. security markets, but whose fundamental risks differ

larger liquidity premia is due to the measure's relative abilities in explaining the cross-section of returns.

from those of individual corporations. This tilts the composition of the cross-section of stocks away from common shares, altering the sample composition. To account for this, we focus on common shares in the remainder of our analysis. This analysis eliminates penny stocks by requiring a minimum daily close price of $1 over the entire previous calendar year. This differs from Amihud (2002) where penny stocks are identified solely based on end-of-year close prices that fall below $5. We obtain qualitatively similar results based on either approach. Table 6 replicates Table 5, focusing on common shares only. Once more, estimated liquidity premia double when we use $OCAM$ in lieu of $CCAM$.

Of note, in the sample of common shares, the declining pattern in liquidity premia manifests itself in liquidity premia that are not significantly different from zero post decimalization. Similar findings have been documented, for example by Ben-Rephael et al. (2013).[24]

Appendix B presents the analogue of Table 6 for a sample that includes both NYSE- and AMEX-listed stocks. Liquidity premia grow by nearly 100% in the more comprehensive sample, indicating that the generally smaller (less liquid) AMEX-listed stocks command larger liquidity premia. Moreover, the $OCAM$ point estimates remain larger than respective estimates for $CCAM$ in this sample. Appendix C extends the analysis to NASDAQ-listed stocks. Liquidity premia for NASDAQ-listed firms are four to five times larger than those for NYSE- and AMEX-listed stocks, and once again liquidity premia based on $OCAM$ are roughly double those based on $CCAM$. Moreover, in Appendix C, we see that our results are robust to estimation using open prices from CRSP for the subset of years for which they are available.

Appendix D shows that the significant liquidity premia found in the sample of *all* NYSE-listed stocks post-decimalization, presented in the last two columns of Table 5, reflect premia

[24]In untabulated results, we find that smaller stocks in the sample still command liquidity premia post decimalization. For example, in the 2001–2017 period, both versions of Amihud's measure have significant explanatory powers for the returns of NYSE-listed stocks whose market-capitalizations fall below the sample median. We estimate monthly liquidity premia of 3bps and 4.2bps based on $CCAM$ and $OCAM$, respectively, in this subsample of small stocks and no premia for stocks with market-capitalizations above the sample median.

Table 6: **Fama-MacBeth estimates of monthly returns on stock characteristics, NYSE-listed common shares, 1964−2017.** This table presents Fama-MacBeth estimates of Equation (7) for NYSE-listed common shares in time periods 1964−2017, 1964−1980, 1981−2000, and 2011−2017. The dependent variable is the monthly stock return in percentage points. $CCAM$ is the traditional Amihud's liquidity measure, and $OCAM$ is Amihud's measure after removing overnight price movements. We divide each $CCAM_{iy}$ and $OCAM_{iy}$ observation by its respective sample mean across stocks in year y, thereby centering each measure to have a mean of one, making coefficients across the two measures comparable. β^{mkt} is market beta estimated across ten size portfolios using daily observations from the last calendar year. $R100$ is the compound return on a stock in the last 100 days of the previous calendar year, and $R100YR$ is the compound return over the remaining trading days in the last calendar year. $\ln(M)$ is the natural log of market capitalization at the end of the previous calendar year. $SDRET$ is the standard deviation of daily returns over the previous calendar year. DYD is the ratio of total cash dividend distribution over the previous calendar year to the closing price at the end of that year, or dividend yield. Newey-West standard errors using two lags are reported in parentheses. Symbols *, **, and *** reflect statistical significance at 10%, 5%, and 1% type one error, respectively.

	1964−2017		1964−1980		1981−2000		2001−2017	
β^{mkt}	0.043	0.059	−0.186	−0.179	0.402	0.401	−0.145	−0.100
	(0.303)	(0.310)	(0.439)	(0.444)	(0.559)	(0.555)	(0.563)	(0.596)
$CCAM$	0.031**		0.052		0.028*		0.014	
	(0.013)		(0.034)		(0.016)		(0.009)	
$OCAM$		0.062***		0.119**		0.053**		0.017
		(0.018)		(0.049)		(0.022)		(0.012)
$R100$	0.502**	0.492**	0.698**	0.678**	0.865***	0.855***	−0.103	−0.104
	(0.214)	(0.213)	(0.341)	(0.342)	(0.272)	(0.270)	(0.474)	(0.472)
$R100YR$	0.289***	0.286***	0.466***	0.467***	0.364**	0.358**	0.030	0.027
	(0.095)	(0.095)	(0.178)	(0.179)	(0.145)	(0.145)	(0.172)	(0.172)
$\ln(M)$	−0.125***	−0.104***	−0.193***	−0.158***	−0.045	−0.018	−0.149***	−0.148***
	(0.034)	(0.034)	(0.065)	(0.060)	(0.062)	(0.064)	(0.047)	(0.049)
$SDRET$	−1.215***	−1.295***	−0.419	−0.541	−1.999***	−2.112***	−1.091	−1.095
	(0.373)	(0.381)	(0.713)	(0.739)	(0.540)	(0.553)	(0.697)	(0.698)
DYD	−0.248	−0.235	−0.678	−0.621	0.354	0.342	−0.519	−0.518
	(0.735)	(0.721)	(2.281)	(2.236)	(0.311)	(0.307)	(0.453)	(0.453)

on non-common shares. This finding is in line with one in Ben-Rephael et al. (2013), who argue that the diminished liquidity premium among common shares in recent years reflects the introduction of alternative investment instruments that facilitate indirect claims to common shares. For example, investors holding ETFs are not exposed to trading costs of the

underlying assets, even though they invest in them.

We now show that $OCAM$ does a better job than $CCAM$ of explaining the cross-section of stock returns, something that need not be implied by the larger liquidity premium associated with $OCAM$. Due to the high correlation between $CCAM$ and $OCAM$, we decompose each version of Amihud's measure into two linearly orthogonal components with respect to the other version.[25] We estimate

$$CCAM_{ity} = \lambda^0_{ty} + \lambda^1_{ty} OCAM_{ity} + Z_{ity} \tag{8}$$

and

$$OCAM_{ity} = \gamma^0_{ty} + \gamma^1_{ty} CCAM_{ity} + \tilde{Z}_{ity} \tag{9}$$

every month, and store the corresponding residuals Z_{ity} and $\tilde{Z}_{ity}$. We then estimate Equation (7) using the Fama-MacBeth approach, replacing either $CCAM$ or $OCAM$ with the corresponding residual version of Amihud's measure, respectively Z_{ity} or $\tilde{Z}_{ity}$. The coefficient on each residual uncovers the incremental informational content of each version of Amihud's measure over its alternative counterpart for the pricing of liquidity.

Table 7: **Fama-MacBeth estimates for pricings of orthogonally-decomposed measures, NYSE-listed common shares, 1964−2017.** This table presents Fama-MacBeth estimates of Equation (7) for NYSE-listed common shares in time periods 1964−2017, 1964−1980, 1981−2000, and 2011−2017. The dependent variable is monthly stock return in percentage points. Independent variables are identical to those in Table 6, save for the measures of stock liquidity. Z is substituted for $CCAM$, and reflects the residuals from Equation (8). $\tilde{Z}$ is substituted for $OCAM$, and reflects the residuals from Equation (9). Newey-West standard errors using two lags are reported in the parentheses. Symbols *, **, and *** reflect statistical significance at 10%, 5%, and 1% type one error, respectively.

	1964−2017		1964−1980		1981−2000		2001−2017	
Z	-0.035^{**}		-0.074^{*}		-0.024		-0.008	
	(0.017)		(0.044)		(0.021)		(0.020)	
$\tilde{Z}$		0.072^{***}		0.144^{**}		0.059^{**}		0.018
		(0.022)		(0.058)		(0.027)		(0.024)

[25] Lou and Shu (2017) and Barardehi et al. (2019) adopt similar approaches.

Table 7 shows that when we remove the information contained in $CCAM$ related to variation in $OCAM$, the residual is not priced. Indeed, it produces a negative coefficient, possibly reflecting pricing of volatility, rather than liquidity. In contrast, the residual from regressing $OCAM$ on $CCAM$, i.e., $\tilde{Z}$, is priced over the entire sample period, as well as in the four last decades of the previous millennium. In essence, little relevant information is lost when we take out the information contained in the classical $CCAM$ Amihud measure. Phrased differently, including the noise in the form of close-to-open returns appears to add measurement error that attenuates estimates of liquidity premia. We further observe that although the "pricing" of each residual loses statistical significance in the 2001-2017 post-decimalization time period, the coefficient signs are consistent with those found in earlier time periods.

Our findings using Fama-MacBeth regressions showcase the meaningful improvements in Amihud's liquidity measure that obtain when one uses open-to-close returns to measure price impacts rather than close-to-close returns. However, this approach is not designed to control for unobserved temporally-fixed stock characteristics when estimating coefficients, and it cannot account for fixed or varying auto-correlations in the error terms that may lead to inflated t-statistics (see Petersen (2009)). Another possible concern with our previous estimates is that we match all monthly returns observations in year y with measures constructed using data from year $y-1$. This matching suggests that investors only care about information from last calendar year rather than more recent information.

We next take a conservative estimation approach, which establishes that our findings are robust to these potential biases. We estimate panels of monthly stock returns that are matched with stock characteristics constructed from the most recent twelve months of data (see Section 2). We use a Fama-French four-factor model to construct risk-adjusted returns as the dependent variable in our panel regressions. Using risk-adjusted returns as dependent variables deals with the "errors-in-variables" issue when test assets are individual stocks (Brennan et al. 1998). Imposing the functional form via which systematic risk factor load-

ing and expected returns are related both shifts measurement error to the left-hand-side and soaks up variation in the dependent variable. This restricted model also reduces the flexibility with which coefficients are estimated. For these reasons the estimation procedure is more conservative than that in Equation (7). In addition, panel regressions let us control for invariant stock and time characteristics by use of stock and month-year fixed effects. Finally, we cluster standard errors at both month-year and stock levels. That is, we estimate

$$\begin{aligned} RA_{it} &= \alpha_0 + \alpha_1 LIQ_{i,t-1} + \alpha_2 RET_{i,t-1}^{t-4} + \alpha_3 RET_{i,t-5}^{t-12} + \alpha_4 \ln(M_{i,t-12}) \\ &+ \alpha_5 SDRET_{i,t-1} + \alpha_6 DYD_{i,t-1} + \text{fixed effects} + u_{it}, \end{aligned} \tag{10}$$

where RA_{it} are risk-adjusted returns from a Fama-French four-factor model; fixed effects include both month-year and stock effects; $LIQ \in \{CCAM, OCAM\}$; and standard errors are clustered at *both* time and firm levels to account for auto-correlations in the error term, u_{it}, following Cameron et al. (2012). For a given liquidity measure, the liquidity premium, in basis points, is the product of the coefficient on the liquidity measure and the measure's sample average.[26]

Table 8 reinforces our earlier findings. Even after controlling for stock fixed effects, and accounting for the auto-correlations in the error terms, we obtain significantly positive liquidity premia using both versions of Amihud's measure. Again, liquidity premia obtained using $OCAM$ are larger than those obtained using $CCAM$ in the entire sample period, as well as in the first and second subsample periods of our data. Temporal patterns are also similar to what we found earlier: the premium on monthly risk-adjusted returns based on $OCAM$ goes from 12.5 basis points per month in 1964-1980 to 5.9 basis points in 1981-2000, before vanishing post-decimalization.

Our final asset pricing tests conduct the panel regression analogues of the orthogonal

[26]The panel analysis controls for temporal changes in all variables of the model by including month-year fixed effects. As such, we do not center liquidity measures to have sample means of 1 to avoid creating inconsistencies with how time fixed effects control for common time variations. Accordingly, liquidity premia must be calculated as the product of the regression coefficient and sample mean.

Table 8: **Panel regression estimates of monthly risk-adjusted returns on stock characteristics, NYSE-listed common shares, 1964−2017.** This table presents GLS estimates of Equation (10) for NYSE-listed common shares in time periods 1964−2017, 1964−1980, 1981−2000, and 2011−2017. The dependent variable is the monthly risk-adjusted returns, in percentage points, based on a Fama-French four-factor model. $CCAM$ is the traditional Amihud's liquidity measure, and $OCAM$ is Amihud's measure after removing overnight price movements. Both measures are constructed monthly, using daily absolute return per dollar observations from the preceding twelve months. Liquidity premium, in basis points, is the product of the coefficient on the liquidity measure and the measure's average in the corresponding sample. Dividend yield, $DYD_{i,t-1}$, divides total dividend distributions between months $t-12$ and $t-1$ by the closing price at the end of month $t-1$. Momentum measures $RET_{i,t-1}^{t-4}$ and $RET_{i,t-5}^{t-12}$, respectively, capture compound returns over the preceding three months and the nine months before that. Return volatility $SDRET_{i,t-1}$ is given by the standard deviation of daily stock returns over the preceding 12 months. Market capitalization, $M_{i,t-12}$, is the product of shares outstanding and the closing price at the end of the month, a year earlier. Both year-month and stock fixed effects are included. Standard errors are clustered at both year-month and stock levels. Symbols *, **, and *** reflect statistical significance at 10%, 5%, and 1% type one error, respectively.

	1964−2017		1964−1980		1981−2000		2001−2017	
$CCAM$	0.151**		0.120		0.264**		0.193	
	(0.063)		(0.078)		(0.120)		(0.540)	
$OCAM$		0.414***		0.406**		0.697**		0.330
		(0.147)		(0.180)		(0.345)		(1.134)
RET_{-1}^{-4}	−0.079***	−0.079***	−0.145***	−0.145***	−0.103***	−0.103***	−0.084***	−0.084***
	(0.011)	(0.011)	(0.017)	(0.017)	(0.014)	(0.014)	(0.024)	(0.024)
RET_{-5}^{-12}	−0.077***	−0.076***	−0.128***	−0.122***	−0.076***	−0.075***	−0.254***	−0.254***
	(0.022)	(0.022)	(0.0340	(0.033)	(0.029)	(0.029)	(0.044)	(0.044)
$\ln(M)$	−0.98***	−0.97***	−1.98***	−1.93***	−1.53***	−1.51***	−1.36***	−1.36***
	(0.05)	(0.05)	(0.14)	(0.14)	(0.12)	(0.12)	(0.14)	(0.14)
$SDRET$	−1.75***	−1.81***	−2.37***	−2.60***	−2.87***	−2.94***	−0.92	−0.92
	(0.32)	(0.32)	(0.58)	(0.58)	(0.48)	(0.49)	(0.56)	(0.56)
DYD	−0.76***	−0.76***	−3.12***	−3.12***	−0.46***	−0.45***	−2.61***	−2.61***
	(0.22)	(0.22)	(0.93)	(0.93)	(0.07)	(0.07)	(0.69)	(0.69)
Observations	793,578		240,321		308,243		244,971	
R^2	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.04
Premium (bps)	3.8	5.5	7.0	12.5	4.3	5.9	0.5	0.7

decomposition analysis presented in Table 7. We fit

$$CCAM_{it} = \theta_0 + \theta_1 OCAM_{it} + U_{it} \tag{11}$$

and

$$OCAM_{it} = \omega_0 + \omega_1 CCAM_{it} + \tilde{U}_{it} \tag{12}$$

using OLS. We then substitute the residuals U_{it} and $\tilde{U}_{it}$ in the second stage for $CCAM_{it}$ and $OCAM_{it}$, respectively, when estimating Equation (10). Table 9 reveals that $OCAM$ not only outperforms $CCAM$ in explaining the cross-section of raw returns, it does so for the the cross-section of risk-adjusted returns, and when we account for stock fixed effects and potential auto-correlations in error terms.

Table 9: **Panel regression estimates for pricings of orthogonally-decomposed measures, NYSE-listed common shares, 1964−2017.** This table presents GLS estimates of Equation (10) for NYSE-listed common shares in time periods 1964−2017, 1964−1980, 1981−2000, and 2011−2017. The dependent variable is the monthly risk-adjusted return, in percentage points, based on a Fama-French four-factor model. Independent variables are identical to those in Table 8, expect for the measures of stock liquidity. U is substituted for $CCAM$, and reflects the residuals from Equation (11). $\tilde{U}$ is substituted for $OCAM$, and reflects the residuals from Equation (12). Both year-month and stock fixed effects are included. Standard errors are clustered at both year-month and stock levels. Symbols *, **, and *** reflect statistical significance at 10%, 5%, and 1% type one error, respectively.

	1964−2017		1964−1980		1981−2000		2001−2017	
U	0.008		−0.154		0.266		−7.321	
	(0.142)		(0.153)		(0.261)		(4.562)	
$\tilde{U}$		1.36***		1.11***		2.17**		15.80
		(0.41)		(0.38)		(0.99)		(9.57)

5 Information-driven overnight returns

We conclude our analysis by documenting the information-driven role of overnight price movements in distorting $CCAM$ from capturing liquidity properly. To do this, we exploit cross-stock variation in the time distance between the last transaction of a typical trading day and official close (4:00pm EST). We establish a very strong association between the extent of non-trading before close and the disparity between $CCAM$ and $OCAM$. Variation in non-trading, which may reflect frictions such as lack of liquidity, large minimum tick size, etc., in-

troduces heterogeneity in the stock of accumulated "overnight" information flow. The econometrician observes the price implications of overnight information flow on a given day as the difference between the close price and the open price of the next day.[27] The extent of information incorporated into prices outside of trading hours is more difficult to observe. We use the time span between the last trade of a trading day and the close as a proxy for the accumulated information that is not incorporated into closing price. Figure 4 illustrates the mapping between the extent of non-trading prior to close on a given trading day and the amount of accumulated information that is expected to be impounded in price the following trading day(s).

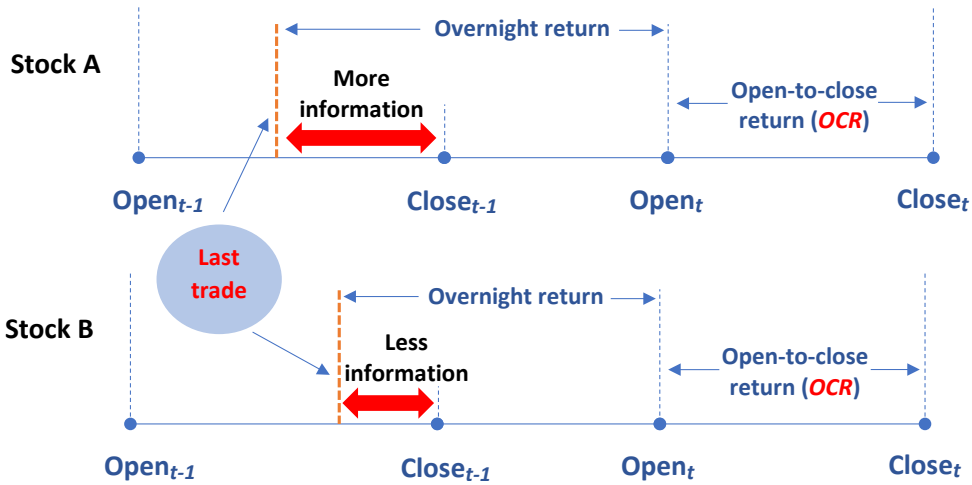

Figure 4: **Illustration of non-synchronous trading and the extent of overnight information accumulation.** The figure illustrates the relationship between the extent of non-trading before close on date $t-1$ and the amount of information contained in overnight returns.

Using intraday transaction data from the TAQ database between 1993−2013, we measure the time distance, in hours, between the last transaction and close each trading day. We then average these distances every year for each stock to construct HTC_{iy} (hours to close),

[27] When there is no closing price available, i.e., when there is not a "closing cross" that corresponds to the final transaction at close, CRSP reports the midpoint of the best bid and ask prices at 4:00pm as the "close price." Importantly, this type of closing price which is not associated with trading may differ from the price associated with the last transaction the same trading day. If anything, when such differences are meaningfully large they introduce noise in our proxy of overnight information accumulation, i.e., the time distance between the last transaction of the day and 4:00pm, attenuating our results. We observe that any empirical analysis that relies closing prices is exposed to this sort of measurement error.

which measures the average extent of non-trading before close in a given year.

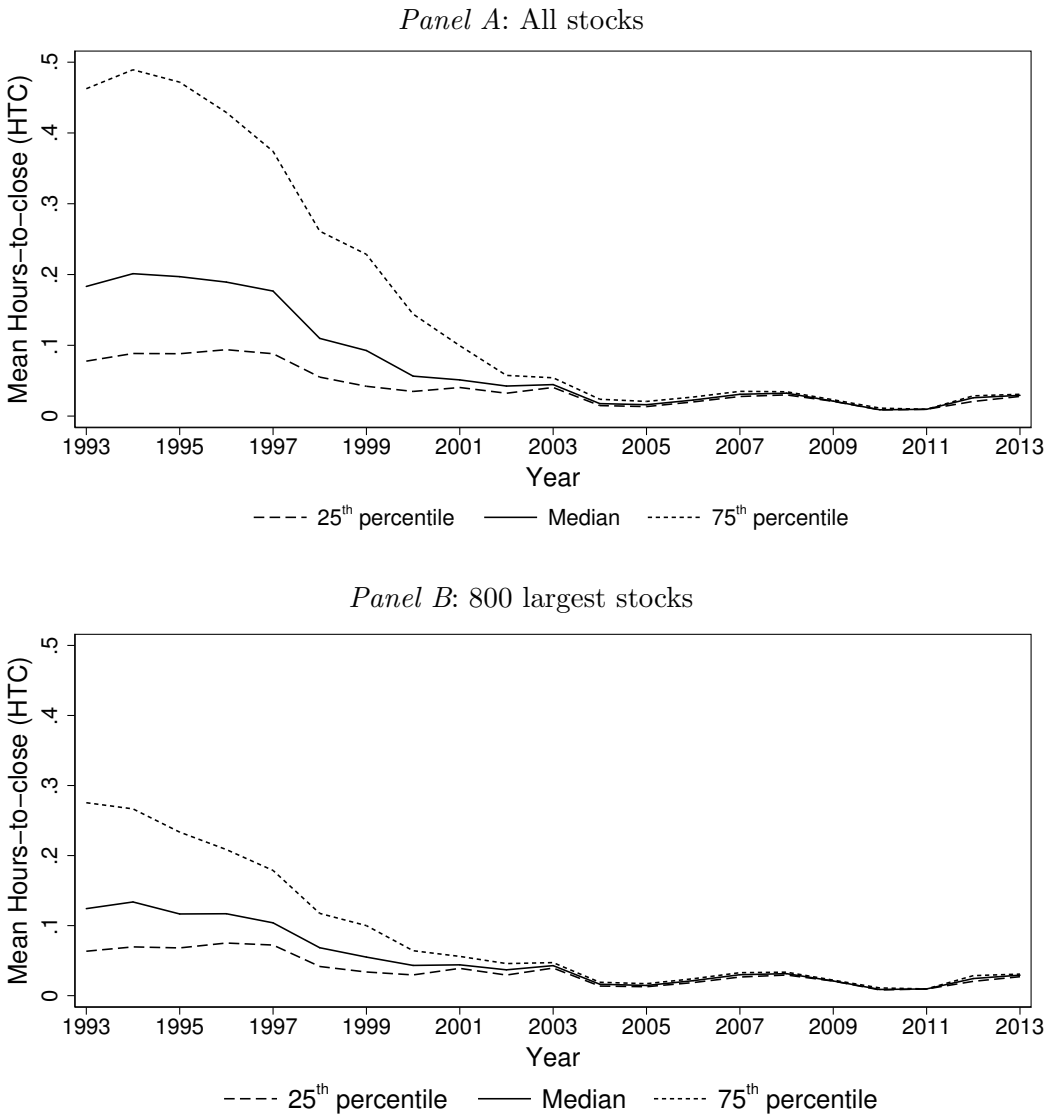

Figure 5: **Evolution of the the extent of non-trading.** This figure plots temporal changes in the cross-stock distribution of HTC in the 1993-2013 period. For each stock i in year y, HTC_{iy} measures the average time distance, in hours, of the last transaction and close per trading day. The year-specific quartiles of HTC are plotted against time. Panel A presents statistics in the sample of all stocks, and Panel B presents the patterns for the sample of largest 800 stocks based on market-capitalizations at the end of the preceding year.

Figure 5 shows that, pre-decimalization, non-trading exhibits remarkable variation in the cross-section. For instance, in 1993, the first and third quartiles of the mean time distance

between the last trade and close are roughly 6 and 28 minutes, respectively—the last transaction was well before close for most stocks. This variation begins to vanish as decimalization is implemented 1997−2001, and disappears after 2003, once automated trading dominates markets. These patterns mirror those for the cross-sectional distribution of ROC in Figure 2, suggesting that variation in non-trading is related to variation in ROC (the extent of disparity between $CCAM$ and $OCAM$). Figure 6 plots ROC_{iy} against HTC_{iy}, underscoring the strong association between the extent of non-trading and differences between $CCAM$ and $OCAM$: the ratio ROC declines with HTC with a slope of roughly $-1/3$. Figures 5 and 6 reveal qualitatively identical patterns when we control for changes in sample composition by focusing on the 800 largest stocks according to market-capitalizations at the end of the previous year.

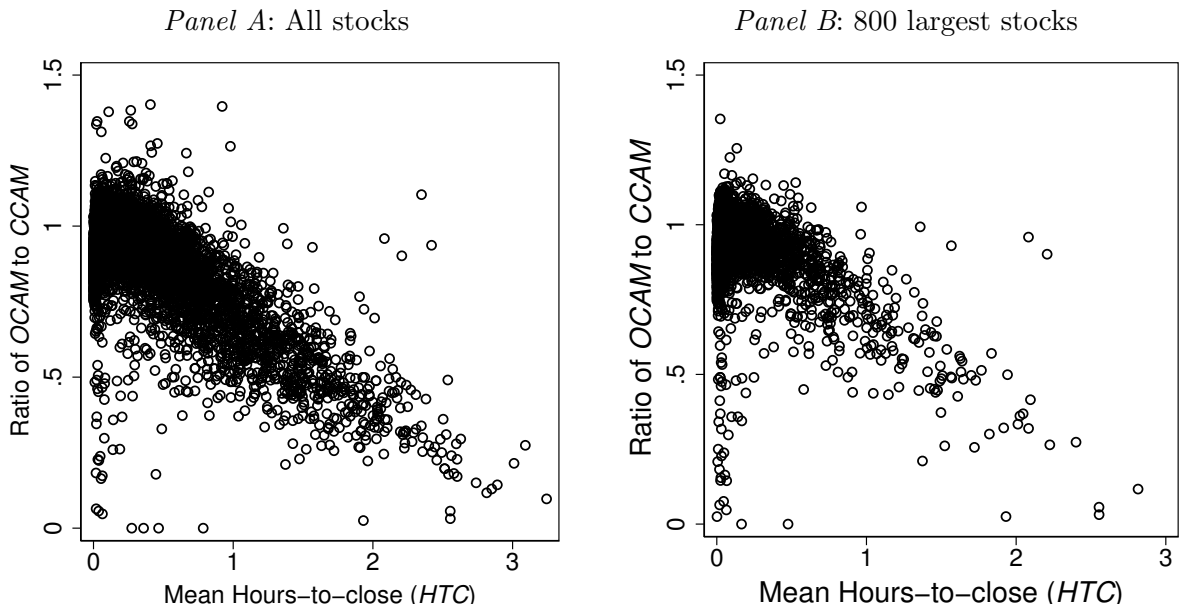

Figure 6: **Association between $OCAM$-to-$CCAM$ ratio and HTC.** The figure presents a scatter plot of ROC_{iy} against HTC_{iy} in the 1993-2013 period. For each stock i in year y, ROC_{iy} is the ratio of open-to-close Amihud measure, $OCAM$, to the traditional Amihud measure, $CCAM$; and HTC_{iy} measures the average time distance, in hours, of the last transaction and close (4:00pm EST) per trading day.

To quantify the ability of HTC to explain variations in the disparity between $CCAM$ and $OCAM$, we add HTC_{iy} to a regression of ROC_{iy} on stock characteristics. The panel

Table 10: **Association between $OCAM$-to-$CCAM$ ratio and stock characteristics, including non-trading.** This table presents panel regression estimates when ROC_{iy} is regressed on the natural logs of market-capitalization, $\ln(M_{i,y-1})$, mean daily turnover, $\ln(TR_{i,y-1})$, and a measure of non-trading, HTC_{iy} in the 1964−2017 sample. For stock i in year y, HTC_{iy} reflects the average time, in hours, between the last transaction and close on a trading day. Specifications differ in inclusion of HTC and in the sets of fixed effects introduced. Numbers in parentheses reflect standard errors of estimates, clustered at both stock and year levels. Symbols *, **, and *** reflect statistical significance at 10%, 5%, and 1% type one error, respectively.

Dep. Var. = ROC								
$\ln(M)$	0.006^{*}	-0.014^{***}	0.024^{***}	-0.016^{***}	0.005	-0.014^{***}	-0.003	-0.013^{***}
	(0.003)	(0.001)	(0.004)	(0.003)	(0.004)	(0.001)	(0.002)	(0.001)
$\ln(TR)$	0.027^{***}	-0.020^{***}	0.060^{***}	-0.025^{***}	0.031^{***}	-0.011^{***}	0.017^{***}	-0.010^{***}
	(0.005)	(0.003)	(0.0060	(0.005)	(0.006)	(0.003)	(0.004)	(0.002)
HTC		-0.300^{***}		-0.274^{***}		-0.307^{***}		-0.282^{***}
		(0.010)		(0.010)		(0.008)		(0.008)
Stock FE	No		Yes		No		Yes	
Year FE	No		No		Yes		Yes	
Adj-R^2	0.05	0.36	0.44	0.63	0.09	0.40	0.50	0.68
Observations	21,446							

regression includes stock fixed effects to account for unknown omitted stock attributes that may explain cross-stock variation in ROC. We also use year fixed effects to account for systematic temporal variation (e.g., varying sample composition) that may explain temporal variation in ROC's year-specific moments. Table 10 shows that HTC strongly explains the variation in the disparity between $CCAM$ and $OCAM$ even *after* controlling for stock characteristics and fixed effects. In fact, in presence of all other controls, the adjusted-R^2 rises by a striking 18 percentage points when we include HTC as an additional explanatory variable. The negative coefficient on HTC indicates that longer periods of non-trading before close are associated with greater disparities between $CCAM$ and $OCAM$. Concretely, a one-hour increase in the average duration of non-trading before close is associated with roughly a 0.3 (30 percentage point) decline in ROC—i.e., a 5 percentage point decline in ROC for each additional 10 minutes of non-trading before close.[28] The estimated slope coefficient is very

[28]Qualitatively similar findings obtain when we use HTC from the previous year to explain variation in

close to the visually discerned slope of $-1/3$ from Figure 6.

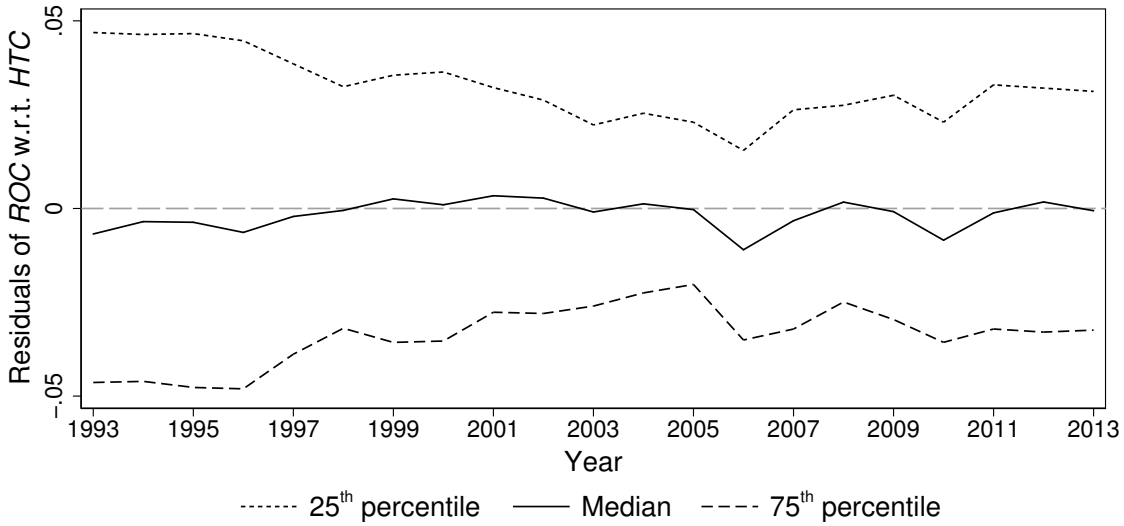

Figure 7: **Evolution of the $OCAM$-to-$CCAM$ ratio after controlling for non-trading.** The figure plots the temporal changes in the cross-stock distribution of the residuals from annual cross-sections of ROC_{iy} on HTC_{iy} in the 1993-2013 period. For each stock i in year y, ROC_{iy} is the ratio of open-to-close Amihud measure, $OCAM$, to the traditional Amihud measure, $CCAM$; and HTC_{iy} measures the average time distance, in hours, of the last transaction and close (4:00pm EST) per trading day. The year-specific quartiles of residuals are plotted against time.

To provide additional evidence of the co-variation between ROC and HTC, we document the temporal evolution of the cross-sectional variation in ROC after accounting for the variation explained by HTC. To do this, for each year, we obtain the residuals from cross-sectional regressions of ROC_{iy} on HTC_{iy}, and then investigate the temporal evolution of the cross-sectional distribution of these residuals. These cross-sectional regressions feature an average R^2 of 59.4% for the 1993−2013 period; and the average R^2s for the sub-samples 1993−2002 and 2003−2013 are 77.6% and 40.7%, respectively, further underscoring the relevance of information-driven price movements for our results. Figure 7 highlights the ability of HTC to explain variation in ROC. In contrast to the large cross-sectional and temporal variations in ROC highlighted in Figure 5, the residuals of ROC on HTC are concentrated around zero (over half of the observations always fall between -0.05 and 0.05) and are remarkably stable

current year's ROC.

over time, indicating that much of ROC's variation is due to variation in the extent of non-trading near close, especially pre-decimalization and before automation of equity markets.[29]

6 Conclusion

Amihud's (2002) liquidity measure ($CCAM$) has been widely used by researchers to study the importance of stock liquidity for an array of financial economics issues, ranging from asset pricing to corporate finance. Its usage, in part, reflects the measure's simple construction using data that can be obtained for long histories and across different markets. The many insights based on this measure make its precision crucial. Our paper develops and implements simple improvements to this measure that require almost no additional data processing effort.

Our $OCAM$ modification uses open-to-close returns, rather than close-to-close returns, to address a time mismatch in the construction of $CCAM$. Our modified measure better explains the cross-section of returns, revealing that liquidity premia are substantially larger than previously understood. We provide strong evidence that $OCAM$ better explains the cross-section of expected (risk-adjusted) returns than $CCAM$. Liquidity premia based on $OCAM$ are double those based on $CCAM$. Including overnight returns in the Amihud measure adds measurement error that sharply attenuates estimates of liquidity premia.

Finally, we exploit cross-sectional and temporal variation in the extent of non-trading before close as a proxy for variation in information-driven price movements to gain insights into the sources of differences between $CCAM$ and $OCAM$. We find that this proxy explains a large share of the cross-sectional variation in these differences. Overall, our paper highlights the importance of excluding information-driven price movements when constructing measures of stock liquidity.

[29] Such stability of residuals does *not* obtain when one regresses ROC on year fixed effects only.

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.

Amihud, Y., A. Hameed, W. Kang, and H. Zhang. 2015. The illiquidity premium: International evidence. *Journal of Financial Economics*, 117(2), 350–368.

Amihud, Y. and J. Noh. 2018. Illiquidity and stock returns – II: Cross-section and time-series effects. Working paper.

Anand, A., P. Irvine, A. Puckett, and K. Venkataraman. 2013. Institutional trading and stock resiliency: Evidence from the 2007-2009 financial crisis. *Journal of Financial Economics* 108, 773-797.

Asparouhova, E., H. Bessembinder, and I. Kalcheva. 2010. Liquidity biases in asset pricing tests. *Journal of Financial Economics* 96, 215–237.

Asparouhova, E., H. Bessembinder, and I. Kalcheva. 2013. Noisy prices and inference regarding returns. *Journal of Finance* 68, 655–714.

Atchison M. D., K. C. Butler and R. R. Simonds. 1987. Nonsynchronous security trading and market index autocorrelation. *Journal of Finance*

Barardehi, Y.H., D. Bernhardt, and R.J. Davies. 2019. Trade-time measures of liquidity. *Review of Financial Studies*. 32(1), 126–179.

Barclay, M. J. and T. Hendershott 2003. Price discovery and trading after hours. *Review of Financial Studies* 16, 1041–73.

Barclay, M. J. and T. Hendershott 2008. A comparison of trading and non-trading mechanisms for price discovery. *Journal of Empirical Finance* 15, 839–849.

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

Brennan, M. J., T. Chordia, and A. Subrahmanyan. 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal*

of Financial Economics 49, 345–73.

Belcourt, T. 1996. Three Essays in Applied Microeconomic Theory. PhD Thesis. *Queen's University.*

Boehmer, E., K. Fong, and L. Wu. 2015. International evidence on algorithmic trading. Working Paper.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2012. Robust Inference with Multiway Clustering. *Journal of Business & Economic Statistics* 29, 238–249.

Cao, C., E. Ghysels and F. Hatheway. 2000. Price discovery without trading: evidence from the Nasdaq pre-opening. *Journal of Finance* 55, 1339–1365.

Chen, Z., Y.Huang, Y. Kusnadi, and K.J. Wei. 2017. The real effect of the initial enforcement of insider trading laws. *Journal of Corporate Finance*, 45, 687–709.

Chordia, T., R. Roll, and A. Subrahmanyan. 2000. Commonality in liquidity. *Journal of Financial Economics* 56, 3–28.

Cliff M., M. J. Cooper, and H. Gulen. 2008. Return difference between trading and non-trading hours: like night and day. Working paper.

Drienko, J., T. Smith, A. Von Reibnitz. 2017. A review of the return-illiquidity relationship, *Critical Finance Review*, forthcoming.

Harris, L. and A. Amato. 2018. Illiquidity and Stock Returns: Cross-Section and Time-Series Effects: A Replication (2002). *Critical Finance Review*, forthcoming.

Hasbrouck J. 2006. Trading costs and returns for U.S. equities: Estimating effective costs from Daily Data. *Journal of Finance* 64, 1445–77.

Hendershott T., D. Livdan, and D. Rösch. 2018. Asset pricing: A tale of night and day. Working paper.

Hung, M., X. Li, and S. Wang. 2014. Post-earnings-announcement drift in global markets: Evidence from an information shock. *Review of Financial Studies*, 28(4), 1242–1283.

Kahle, K.M. and R.M. Stulz. 2017. Is the U.S. Public Corporation in Trouble?. *Journal of Economic Perspectives*, 31(3), 67–88.

Kingsley Y.L.F., C.W. Holden and C.A. Trzcinka. 2017. What Are the Best Liquidity Proxies for Global Research?. *Review of Finance*, 21(4), 1355–1401.

Lang, M. and L. Stice-Lawrence. 2015. Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics*, 60(2-3), 110-135.

Lee, K. H. 2011. The world price of liquidity risk. *Journal of Financial Economics*, 99(1), 136–161.

Lipson M. L. and S. Mortal. 2009. Liquidity and capital structure. *Journal of Financial Markets* 12, 611–644.

Lou, X. and T. Shu. 2017. Price impact or trading volume: Why is the Amihud (2002) measure priced? *Review of Financial Studies* 30, 4481–4520.

Petersen, M.A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22, 435–480.

Sadka, R. 2006. Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. *Journal of Financial Economics* 80, 309–349.

Santosh S. 2016. The speed of price discovery: trade time vs clock time. Working paper.

Stoll, H. and R. Whaley. 1990. Stock market structure and volatility. *Review of Financial Studies* 3, 37–71.

A $OCAM$-$CCAM$ disparity and price impacts of the open auction

In this section, we address the possibility that $OCAM$ might fail to align returns and trading volumes by including the simultaneously-determined price and volume associated with the "open auction."[30] We take a conservative approach, constructing three alternative versions of $OCAM$ that exclude price and trading volume in the first x minutes of trading, with $x \in \{5, 10, 15\}$, using trade information from the remainder of the trading day as inputs for our open-to-close Amihud measure of liquidity. For each trading day, instead of open price, we use the price of the last transaction executed at most x minutes after 9:30am, only using the open price if there is no transaction in that window. Accordingly, we subtract trading volume realized in those x minutes from the aggregate daily volume. We then calculate corresponding "open-to-close" returns and dollar volumes from the remainder of the trading day to construct the daily proxy of price impacts, i.e., per dollar absolute open-to-close return. By averaging this proxy annually for each stock and each $x \in \{5, 10, 15\}$, we construct alternative $OCAM^x_{iy}$ measures that exclude most open auction trades.[31]

Figure A.8 shows that trading dynamics within the first minutes of the trading day, which includes the opening auction, minimally impact magnitudes of $OCAM$. The correlation coefficients between the natural log of $OCAM$ and the alternative $OCAM^x$ measures always exceed 96% (exceeding 98% except in the first two years of the sample), with an average correlation of over 99%.[32] As such, exclusion of trade and price data that may overlap with

[30] The TAQ database does not identify opening trades for most market centers, including the NYSE.

[31] We otherwise follow the construction of $OCAM_{iy}$ detailed in Section 2.1. Of note, we can only construct these versions of $OCAM$ post 1992, i.e., only once TAQ data became available.

[32] Average correlations between $OCAM$ amd $OCAM^x$, before taking logs, exceed 95%. Taking natural logs

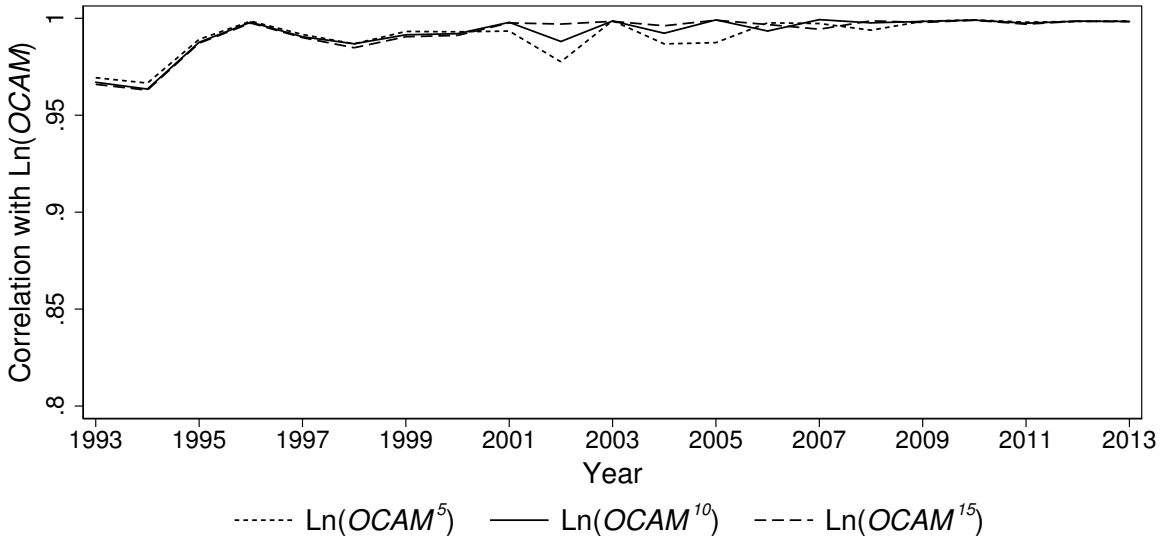

Figure A.8: **Cross-sectional correlations between** $\ln(OCAM)$ **and** $\ln(OCAM^x)$ **by year.** The figure plots the cross-stock correlation coefficients between $\ln(OCAM)$ and $\ln(OCAM^x)$, with $x \in \{5, 10, 15\}$ by year. $OCAM^x$ is an alternative version of $OCAM$ where open price is substituted by the last transaction price within x minutes after 9:30am on the same day, and measures of daily volume that exclude trading volume realized within the first x minutes of trading. Cross-stock correlations of $OCAM$ versus each of its three alternative versions are calculated annually in the 1993–2013 period.

opening trades appear to have negligible effects on $OCAM$ quantities, revealing that the open auction does not drive variation in our measure.

B Liquidity premia on NYSE- and AMEX-listed common shares 1964-2017.

This section extends the analysis of Amihud (2002) for the 1964−2017 period, expanding the sample to include both NYSE- and AMEX-listed common shares. Table B.8 reinforces our findings presented in Table 6. Liquidity premia based on $OCAM$ are approximately twice

of these annual measures addresses non-linearities at the distribution tails. Such non-linearities likely reflect data errors that give rise to a few extreme price/volume observations at the daily level. For instance, in a few cases, we excluded from the sample trading days with reported trading volume in the first x minutes that *exceeded* the daily trading volume reported by CRSP. Other data errors are not as easy to identify and exclude.

those based on $CCAM$. Adding AMEX-listed stocks, which are considered to be less liquid than NYSE-listed stocks, results in liquidity premia for the sample of NYSE- and AMEX-listed firms that are nearly double those for the sample of NYSE-listed stocks. Moreover, Table B.8 shows that weighting observations using the previous month's gross returns, as proposed by Asparohouva et al. (2010), only leads to marginal declines in liquidity premia estimates, leaving results qualitatively unchanged.

C Liquidity premia on NASDAQ-listed common shares 1993-2017.

This section extends the analysis of Amihud (2002) to NASDAQ-listed common shares for the 1993−2017 period, contrasting outcomes with those for NYSE- and AMEX-listed common shares using CRSP for opening prices rather than GFD. Table C.8 performs Fama-MacBeth regressions similar to Table 5, reinforcing our earlier findings. In particular, once more, liquidity premia based on $OCAM$ are roughly double those based on $CCAM$. Moreover, liquidity premia for NASDAQ-listed firms are four to five times larger than those for NYSE and AMEX-listed stocks. Further, for this sample period, opening prices are available on CRSP, and the analysis confirms our findings regardless of whether we use CRSP or GFD for opening prices. This underscores the relevance of our analysis for most researchers who will be using publicly available data. Furthermore, it provides evidence that we can trust our findings using GFD prices for earlier dates for which other sources of opening prices do not currently exist. Finally, Table C.8 establishes robustness of our findings to upward biases in point estimates due to noisy prices. The qualitative differences in liquidity premia based

Table B.8: **Fama-MacBeth estimates of monthly returns on stock characteristics, NYSE- and AMEX-listed common shares, 1964−2017.** This table presents Fama-MacBeth estimates of Equation (7) for NYSE- and AMEX-listed common shares in time periods 1964−2017, 1964−1980, 1981−2000, and 2011−2017. The dependent variable is the monthly stock return in percentage points. $CCAM$ is the traditional Amihud's liquidity measure, and $OCAM$ is Amihud's measure after removing overnight price movements. We divide each $CCAM_{iy}$ and $OCAM_{iy}$ observation by its respective sample mean across stocks in year y, thereby centering each measure to have a mean of one. β^{mkt} is market beta estimated across ten size portfolios using daily observations from the last calendar year. $R100$ is the compound return on a stock in the last 100 days of the previous calendar year, and $R100YR$ is the compound return over the remaining trading days in the last calendar year. $\ln(M)$ is the natural log of market capitalization at the end of the previous calendar year. $SDRET$ is the standard deviation of daily returns over the previous calendar year. DYD is the ratio of total cash dividend distribution over the previous calendar year to the closing price at the end of that year, or dividend yield. Newey-West standard errors using two lags are reported in parentheses. Symbols *, **, and *** reflect statistical significance at 10%, 5%, and 1% type one error, respectively.

	1964−2017		1964−1980		1981−2000		2001−2017	
β^{mkt}	0.493	0.665	0.068	0.347	0.658	0.819	0.724	0.802
	(0.284)	(0.279)	(0.415)	(0.4150	(0.613)	(0.599)	(0.358)	(0.360)
$CCAM$	0.074***		0.106**		0.095***		0.015	
	(0.017)		(0.044)		(0.021)		(0.015)	
$OCAM$		0.115***		0.202***		0.118***		0.025
		(0.023)		(0.060)		(0.027)		(0.022)
$R100$	0.438**	0.435**	0.617*	0.615*	0.793***	0.790***	−0.158	−0.164
	(0.206)	(0.205)	(0.336)	(0.3370	(0.248)	(0.248)	(0.465)	(0.464)
$R100YR$	0.229***	0.237***	0.316**	0.340**	0.384***	0.388***	−0.040	−0.042
	(0.080)	(0.080)	(0.136)	(0.138)	(0.124)	(0.124)	(0.154)	(0.155)
$\ln(M)$	−0.114***	−0.118***	−0.194***	−0.192***	−0.052	−0.068	−0.107***	−0.102***
	(0.029)	(0.027)	(0.054)	(0.048)	(0.050)	(0.048)	(0.039)	(0.040)
$SDRET$	−1.420***	−1.600***	−0.467	−0.843	−2.287***	−2.427***	−1.355**	−1.383**
	(0.331)	(0.336)	(0.618)	(0.632)	(0.487)	(0.497)	(0.618)	(0.624)
DYD	0.587	0.494	2.022	1.720	0.192	0.189	-0.385	-0.374
	(0.625)	(0.606)	(1.899)	(1.839)	(0.310)	(0.306)	(0.415)	(0.421)

on $CCAM$ and $OCAM$ are unaffected when we perform WLS estimation with the previous month's gross returns serving as weights, following Asparouhova et al. (2010).

Table B.8: **Fama-MacBeth WLS estimates of monthly returns on stock characteristics, NYSE- and AMEX-listed common shares, 1964−2017.** This table presents Fama-MacBeth estimates of Equation (7) for NYSE- and AMEX-listed common shares in time periods 1964−2017, 1964−1980, 1981−2000, and 2011−2017. Following Asparouhova et al. (2010), observations are weighted using previous month's gross (one plus) returns. The dependent variable is the monthly stock return in percentage points. $CCAM$ is the traditional Amihud's liquidity measure, and $OCAM$ is Amihud's measure after removing overnight price movements. We divide each $CCAM_{iy}$ and $OCAM_{iy}$ observation by its respective sample mean across stocks in year y, thereby centering each measure to have a mean of one. β^{mkt} is market beta estimated across ten size portfolios using daily observations from the last calendar year. $R100$ is the compound return on a stock in the last 100 days of the previous calendar year, and $R100YR$ is the compound return over the remaining trading days in the last calendar year. $\ln(M)$ is the natural log of market capitalization at the end of the previous calendar year. $SDRET$ is the standard deviation of daily returns over the previous calendar year. DYD is the ratio of total cash dividend distribution over the previous calendar year to the closing price at the end of that year, or dividend yield. Newey-West standard errors using two lags are reported in parentheses. Symbols *, **, and *** reflect statistical significance at 10%, 5%, and 1% type one error, respectively.

	1964−2017		1964−1980		1981−2000		2001−2017	
β^{mkt}	0.591**	0.737***	0.245	0.484	0.752	0.879	0.745**	0.821**
	(0.281)	(0.277)	(0.425)	(0.426)	(0.602)	(0.583)	(0.355)	(0.364)
$CCAM$	0.064***		0.091**		0.081***		0.016	
	(0.017)		(0.043)		(0.022)		(0.016)	
$OCAM$		0.098***		0.172***		0.098***		0.026
		(0.023)		(0.058)		(0.028)		(0.023)
$R100$	0.536***	0.532***	0.717**	0.713**	0.928***	0.922***	−0.104	−0.109
	(0.202)	(0.202)	(0.330)	(0.331)	(0.242)	(0.241)	(0.457)	(0.456)
$R100YR$	0.271***	0.278***	0.357**	0.379***	0.455***	0.458***	−0.031	−0.034
	(0.080)	(0.081)	(0.139)	(0.141)	(0.126)	(0.125)	(0.151)	(0.152)
$\ln(M)$	−0.104***	−0.108***	−0.178***	−0.179***	−0.046	−0.060	−0.099**	−0.093**
	(0.029)	(0.027)	(0.054)	(0.049)	(0.050)	(0.048)	(0.041)	(0.042)
$SDRET$	−1.438***	−1.597***	−0.560	−0.897	−2.332***	−2.448***	−1.259**	−1.291**
	(0.328)	(0.333)	(0.611)	(0.624)	(0.484)	(0.494)	(0.611)	(0.616)
DYD	0.745	0.661	2.444	2.174	0.264	0.252	-0.380	-0.363
	(0.620)	(0.601)	(1.885)	(1.824)	(0.296)	(0.291)	(0.418)	(0.423)

D Liquidity premia on non-common shares 1964-2017.

This section extends the analysis of Amihud (2002) for the 1964−2017 period, limiting the sample to NYSE-listed stocks that are not considered common shares. We estimate Equa-

Table C.8: **Fama-MacBeth estimates of monthly returns on stock characteristics, NYSE- and AMEX-listed versus NASDAQ-listed common shares, 1993−2017 (CRSP data only).** This table contrasts Fama-MacBeth estimates of Equation (7) for NYSE- and AMEX-listed common shares in the 1993−2017 period versus those for NASDAQ-listed counterparts. The dependent variable is the monthly stock return in percentage points. $CCAM$ is the traditional Amihud's liquidity measure, and $OCAM$ is Amihud's measure after removing overnight price movements. We divide each $CCAM_{iy}$ and $OCAM_{iy}$ observation by its respective sample mean across stocks in year y, thereby centering each measure to have a mean of one. β^{mkt} is market beta estimated across ten size portfolios using daily observations from the last calendar year. $R100$ is the compound return on a stock in the last 100 days of the previous calendar year, and $R100YR$ is the compound return over the remaining trading days in the last calendar year. $\ln(M)$ is the natural log of market capitalization at the end of the previous calendar year. $SDRET$ is the standard deviation of daily returns over the previous calendar year. DYD is the ratio of total cash dividend distribution over the previous calendar year to the closing price at the end of that year, or dividend yield. Newey-West standard errors using two lags are reported in parentheses. Symbols *, **, and *** reflect statistical significance at 10%, 5%, and 1% type one error, respectively. Estimates are carried out using both ordinary least squares and weighted least squares, with lagged monthly gross (one plus) return used as weights, to correct for biases identified by Asparouhova et al. (2010).

	OLS				WLS			
	NYSE & AMEX		NASDAQ		NYSE & AMEX		NASDAQ	
β^{mkt}	0.701	0.858*	−0.253	−0.112	0.682	0.809	−0.318	−0.175
	(0.442)	(0.427)	(0.298)	(0.290)	(0.435)	(0.424)	(0.292)	(0.286)
$CCAM$	0.023		0.085**		0.017		0.073**	
	(0.014)		(0.032)		(0.014)		(0.033)	
$OCAM$		0.048**		0.178***		0.037*		0.160***
		(0.020)		(0.044)		(0.020)		(0.044)
$R100$	−0.006	−0.012	−0.021	−0.031	0.078	0.076	0.034	0.026
	(0.329)	(0.327)	(0.242)	(0.243)	(0.326)	(0.323)	(0.234)	(0.235)
$R100YR$	0.064	0.061	0.010	0.011	0.098	0.095	0.015	0.016
	(0.123)	(0.123)	(0.042)	(0.042)	(0.124)	(0.123)	(0.041)	(0.041)
$\ln(M)$	−0.059	−0.053	−0.011	0.006	−0.060	−0.054	0.002	0.016
	(0.040)	(0.040)	(0.045)	(0.046)	(0.040)	(0.039)	(0.045)	(0.046)
$SDRET$	−1.164**	−1.252**	−1.092**	−1.265**	−1.172**	−1.241**	−1.227**	−1.391***
	(0.492)	(0.500)	(0.434)	(0.451)	(0.487)	(0.495)	(0.434)	(0.451)
DYD	−0.148	−0.155	−0.015	0.024	−0.174	−0.177	−0.169	−0.134
	(0.323)	(0.325)	(0.615)	(0.620)	(0.324)	(0.327)	(0.597)	(0.602)

tion (7) only including stocks whose share codes differ from 10 and 11 in the CRSP data base. Table D.8 shows that significant liquidity premia among non-common share stocks obtain in the post-decimalization time period. This analysis shows the reason we find sig-

nificant liquidity premia across all NYSE-listed stocks in the 2001−2017 period (Table 5), but not when we focus on common shares (Table 6). Additionally, our findings using non-common share stocks demonstrate the robustness of our findings that liquidity premia based on $OCAM$ are significantly larger that those found based on $CCAM$. Underestimates of liquidity premia by $CCAM$ do not just reflect a quality that is specific to common shares.

Table D.8: **Fama-MacBeth estimates of monthly returns on stock characteristics, NYSE-listed non-common shares, 1964−2017.** This table presents Fama-MacBeth estimates of Equation (7) for NYSE-listed stocks with CRSP share codes that differ from 10 and 11 in time periods 1964−2017, 1964−1980, 1981−2000, and 2011−2017. The dependent variable is the monthly stock return in percentage points. $CCAM$ is the traditional Amihud's liquidity measure, and $OCAM$ is Amihud's measure after removing overnight price movements. We divide each $CCAM_{iy}$ and $OCAM_{iy}$ observation by its respective sample mean across stocks in year y, thereby centering each measure to have a mean of one. β^{mkt} is market beta estimated across ten size portfolios using daily observations from the last calendar year. $R100$ is the compound return on a stock in the last 100 days of the previous calendar year, and $R100YR$ is the compound return over the remaining trading days in the last calendar year. $\ln(M)$ is the natural log of market capitalization at the end of the previous calendar year. $SDRET$ is the standard deviation of daily returns over the previous calendar year. DYD is the ratio of total cash dividend distribution over the previous calendar year to the closing price at the end of that year, or dividend yield. Newey-West standard errors using two lags are reported in parentheses. Symbols *, **, and *** reflect statistical significance at 10%, 5%, and 1% type one error, respectively.

	1964−2017		1964−1980		1981−2000		2001−2017	
β^{mkt}	0.276	0.184	0.341	−0.498	0.460	0.881	0.003	0.051
	(0.620)	(0.629)	(1.700)	(1.755)	(0.848)	(0.816)	(0.265)	(0.264)
$CCAM$	0.089**		0.220**		0.038		0.020**	
	(0.036)		(0.108)		(0.028)		(0.010)	
$OCAM$		0.133**		0.336**		0.042		0.039**
		(0.052)		(0.154)		(0.046)		(0.015)
$R100$	0.810**	0.754**	0.711	0.575	1.086**	1.065**	0.590*	0.571*
	(0.326)	(0.322)	(0.806)	(0.794)	(0.4990	(0.493)	(0.302)	(0.301)
$R100YR$	0.616***	0.716***	1.250**	1.588**	0.407	0.391	0.238	0.239
	(0.232)	(0.250)	(0.610)	(0.672)	(0.280)	(0.282)	(0.254)	(0.254)
$\ln(M)$	0.034	0.009	0.103	0.047	0.007	−0.017	−0.002	0.002
	(0.048)	(0.046)	(0.1320	0.125	(0.063)	(0.061)	(0.034)	(0.034)
$SDRET$	−0.860	−0.983*	−0.604	−0.953	−1.615*	−1.643*	−0.246	−0.259
	(0.548)	(0.560)	(1.136)	(1.187)	(0.906)	(0.913)	(0.775)	(0.776)
DYD	−1.766	−2.365**	−2.835	−4.797	−0.729	−0.706	−1.914**	−1.901**
	(1.168)	(1.196)	(3.459)	(3.551)	(0.857)	(0.827)	(0.997)	(1.009)

E Corrected Amihud's measure and alternative high-frequency measures of liquidity

Here, we show that the improvements in pricing of Amihud's measure due to removing overnight price movements result in more precise measures of liquidity in the cross-section. To do this, we examine the pair-wise correlations of both $CCAM$ and $OCAM$ vis à vis other commonly-used measures of liquidity as benchmarks. We examine cross-stock correlations at annual and monthly frequencies depending on the frequency at which the alternative measures are available.

We use measures of effective trading costs developed by Hasbrouck (2009), namely, two versions of moments estimates of costs and Gibbs estimates of costs, to investigate how closely the $CCAM_{iy}$ and $OCAM_{iy}$ measures correspond with measures of trading costs at annual frequencies in the $1964-2003$ period.[33] Each year, we calculate the cross-stock correlations between each version of Amihud's measure and the cost estimates, and then report the average correlation coefficient across all years. We perform an analogous analysis for the monthly measures $CCAM_{it}$ and $OCAM_{it}$ to examine them against trade-time measures of liquidity (BBD_{it} and $WBBD_{it}$), size-weighted relative quoted spreads (PSP_{it}), and estimates of Kyle's lambda ($LAMBDA$) for the period $2001-2014$.[34] Table E.8 shows that $OCAM$ displays notably higher correlations with all alternative liquidity measures than $CCAM$, reinforcing that $OCAM$ is a more accurate measure of stock liquidity.

[33] Annual data for this time period are available for download at Professor Joel Hasbrouck's website: `http://people.stern.nyu.edu/jhasbrou/Research/GibbsEstimates2005/`.

[34] The monthly measures are obtained from data used in Barardehi et al. (2019).

Table E.8: $CCAM$ **and** $OCAM$ **vs. other standard measures of stock liquidity.** This table presents average cross-sectional correlations between each $CCAM$ and $OCAM$ vis à vis other high-frequency measures of stock liquidity. **Panel A** presents correlations against annual measures of effective costs calculated by Hasbrouck (2006). $cMdmLog_{iy}$ and $cMdmLogz_{iy}$ are two version of Roll's measure of daily return auto-correlations, respectively, reflecting whether missing daily return observations are dropped or replace by zero. $cLogMean_{iy}$ reflects Gibbs estimates using a market-factor model applied to CRSP closing prices and dividends. Every year, correlations between each measure with $CCAM_{iy}$ or $OCAM_{iy}$ are calculated in the 1964−2003 period, and the averages are reported. **Panel B** presents correlations against monthly measures of stock liquidity used in Barardehi et al. (2019). BBD_{it} and $WBBD_{it}$ measure monthly simple and volume-weighted averages of per-dollar price impacts associated with trad-times of fixed-dollar values. PSP_{it} is the monthly time-weighted average of bid-ask spread-over-mid-points. $LMABDA_{it}$ is the monthly estimates of Kyle's λ. Every month, correlations between each measure with $CCAM_{it}$ or $OCAM_{it}$ are calculated in the 2001−2012 period, and the averages are reported.

Correlations vs. Hasbrouk's estimates of effective costs 1964-2003				
Time period	Measure	$cMdmLog$	$cMdmLogz$	$cLogMean$
1964−2003	$CCAM$	0.39	0.38	0.53
	$OCAM$	0.53	0.49	0.66
1964−1980	$CCAM$	0.35	0.37	0.55
	$OCAM$	0.49	0.46	0.67
1981−2003	$CCAM$	0.42	0.39	0.51
	$OCAM$	0.55	0.51	0.65

Correlations vs. Barardehi et al.'s measures 2001−2014				
Measure	BBD	$WBBD$	PSP	$LAMBDA$
$CCAM$	0.77	0.76	0.72	0.71
$OCAM$	0.82	0.81	0.77	0.75