A simple approximation of intraday spreads using daily data[☆]

Kee H. Chung[a,*], Hao Zhang[b,1]

[a]*School of Management, State University of New York (SUNY) at Buffalo*
Buffalo, NY 14260, United States
College of Business Administration, Chung-Ang University, Seoul, Korea

[b]*E. Philip Saunders College of Business, Rochester Institute of Technology*
Rochester, NY 14623, United States

Abstract

This study examines the relation between the bid-ask spread from the daily CRSP data and the bid-ask spread from the intraday TAQ data. We show that the CRSP-based spread is highly correlated with the TAQ-based spread across stocks using data from 1993 through 2009. The simple CRSP-based spread provides a better approximation of the TAQ-based spread than all other low-frequency liquidity measures in cross-sectional settings. However, the CRSP-based spread is highly correlated with the TAQ spread in time-series settings only for NASDAQ stocks. Overall, our results suggest that the simple CRSP-based spread could be used in lieu of the TAQ-based spread in academic research that focuses on cross-sectional analysis.

JEL classification: G12; G20; G30

Keywords: Bid-ask spreads; TAQ; CRSP; Market liquidity; Information asymmetry; Low-frequency liquidity measures

[☆]The paper benefitted greatly from the valuable comments and suggestions provided by Charles M. Jones, Tarun Chordia, and two anonymous referees. The authors thank Yang Cao (CRSP), Suchismita Mishra, Ted Moorman, Andriy Shkilko, Barry Simorangkir (CRSP), Brian Thomas (CRSP), Qiang Wu, Yuxing Yan (WRDS), Mao Ye, and session participants at CRSP Forum 2010, the 2011 Annual Meeting of the Midwest Finance Association, and the 2012 Financial Management Association International (FMA) Annual Meeting for valuable comments and discussion.

[*]Corresponding author. Tel.: +1 716 645 3262; fax: +1 716 645 3823.
E-mail addresses: keechung@buffalo.edu (K. Chung), hzhang@saunders.rit.edu (H. Zhang).
[1]Tel.: +1 585 475 3981; fax: +1 585 475 6920

1. Introduction

The bid-ask spread is a measure of stock market liquidity that has been frequently employed in market microstructure studies. For example, previous studies (e.g., Christie and Schultz, 1994; Huang and Stoll, 1996; Bessembinder, 2003a) use the bid-ask spread to perform inter-market comparisons of trading costs. In addition, market regulators implement various rules and regulations to reduce the cost of trading, and subsequently assess the efficacy of these rules and regulations by analyzing their impact on the bid-ask spread. For instance, the U.S. Securities and Exchange Commission (SEC) reduced the minimum price variation (i.e., tick size) from $\$1/8^{th}$ to $\$1/16^{th}$ in 1997 and again from $\$1/16^{th}$ to one cent in 2001 with the specific purpose of reducing the bid-ask spread in the U.S. securities markets. Similarly, the SEC enacted the Limit Order Display Rule in an effort to reduce the inside spread of NASDAQ-listed stocks.[1]

Most prior market microstructure research relied on the Trade and Quote (TAQ) data provided by the New York Stock Exchange (NYSE), which involves a tedious process of data downloading, error filtering, and variable calculation. In this study, we propose an alternative method of calculating the bid-ask spread that requires only daily data and minimal computational efforts. Our simple liquidity measure would be useful to those who do not have an access to the TAQ database and/or those who want to incorporate stock market liquidity in their research without having to go through the process that is required for the TAQ database. Our simple liquidity measure is readily available for NYSE/AMEX/NASDAQ stocks from 1993 onwards, which coincides with the time period covered by the TAQ database. Another possible advantage of our simple liquidity measure is its availability beyond the time period covered by the TAQ

[1] See Barclay, Christie, Harris, Kandel, and Schultz (1999) for empirical evidence.

database: one can easily obtain our low-frequency bid-ask spread measure from 1925 to 1942 for all NYSE/AMEX stocks and from 1982 to 1992 for most NASDAQ stocks.[2]

Despite its usefulness as a measure of stock market liquidity and information asymmetry, the usage of the TAQ-based bid-ask spread in other research areas has been limited due, at least in part, to data availability problems. Our study is mainly motivated by the need for readily available liquidity measures in other research areas, such as corporate finance, financial accounting, and asset pricing. Researchers in these areas show that liquidity plays an important role in many financial decisions and the pricing of assets.

For example, prior research underscores possible interactions between stock market liquidity and (1) capital structure (Frieder and Martell, 2006; Lipson and Mortal, 2009), (2) dividend payout and stock repurchase decisions (Banerjee, Gatchev, and Spindt, 2007; Brockman, Howe, and Mortal, 2008), (3) ownership structure (Heflin and Shaw, 2000; Sarin, Shastri, and Shastri, 2000; Brockman, Chung, and Yan, 2009), (4) firm value (Fang, Noe, and Tice, 2009), (5) corporate governance (Chung, Elder, and Kim, 2010), (6) executive compensation (Jayaraman and Milbourn, 2012), (7) corporate innovation (Fang, Tian, and Tice, 2013), (8) institutional investors' stock selection decisions (Falkenstein, 1996; Chung and Zhang, 2011; Huang, 2013), and (9) asset pricing (Amihud and Mendelson, 1986; Spiegel and Wang, 2005).[3] Readily available liquidity measures would be useful to researchers in these areas, especially when they need liquidity measures for a large cross-section of firms in the post-1993 period.

[2] Our simple liquidity measure could be useful to researchers who are interested in the analysis of the 1929 stock market crash.

[3] The bid-ask spread is also widely used as a proxy for market liquidity and information asymmetry in accounting literature (e.g., Greenstein and Semi, 1994; Kim and Verrecchia, 1994; Welker, 1995; Coller and Yohn, 1997; Healy, Hutton, and Palepu, 1999; Leuz and Verrecchia, 2000; Leuz, 2003; Guo, Lev, and Zhou, 2004; Mohd, 2005; Chang, Chen, Liao, and Mishra, 2006; Ali, Chen, and Radhakrishnan, 2007; Jayaraman, 2008; Bushee, Core, Guay, and Hamm, 2010; and Cheng, Dhaliwal, and Neamtiu, 2011).

Previous studies have developed low-frequency liquidity measures using daily *closing prices* from the Center for Research in Security Prices (CRSP). They include Roll (1984), Lesmond, Ogden, and Trizinka (1999), Hasbrouck (2009), and Holden (2009). Goyenko, Holden, and Trzcinka (2009) compare the TAQ-based effective spread with various low-frequency liquidity measures using a sample of 400 randomly selected stocks over the period from 1993 through 2005. They show that the TAQ-based effective spread is highly correlated with these low-frequency liquidity measures.

Other studies estimate stock liquidity using *Ask or High Price* and *Bid or Low Price* in the CRSP database. *Ask or High Price* is the highest trading price during the day or the closing ask price on days when the closing price is not available. Likewise, *Bid or Low Price* is the lowest trading price during the day or the closing bid price on days when the closing price is not available. Eckbo and Norli (2002) and Holden (2009) use these variables on no-trade days to estimate the monthly bid-ask spread of stocks that have at least one no-trade day in a given month. However, as Holden (2009) shows, only 26% of the 62,100 stock-months contain one or more no-trade days during 1993-2005. Corwin and Schultz (2012) develop a bid-ask spread estimator using the CRSP's *Ask or High Price* and *Bid or Low Price*. Using a sample of U.S. stocks from 1993 through 2006, Corwin and Schultz (2012) show that their spread estimates are highly correlated with the TAQ-based effective spread–the average cross-sectional correlation coefficient between the two variables is 0.930 during 1993-1996 and 0.732 during 2001-2006.

In the present study, we propose a simple bid-ask spread measure that can be calculated using the daily data from the CRSP. In contrast to previous studies that use the CRSP's *Ask or High Price* and *Bid or Low Price* to obtain low-frequency liquidity measures, we use data in the

two new fields (i.e., *Ask* and *Bid*) that were added to the CRSP database in December 2005.[4] To our best knowledge, this study is the first to analyze the usefulness of these variables in academic research. In addition, our study differs from previous studies (e.g., Lesmond, Ogden, and Trizinka, 1999; Corwin and Schultz, 2012) in that our low-frequency liquidity measure requires neither a sophisticated estimation procedure nor large computational efforts (in both time and space), making it easier for both researchers and practitioners to use.

We show that the CRSP-based spread is highly correlated with the TAQ-based spread using data from 1993 through 2009. For instance, we find that the annual average of monthly cross-sectional correlation coefficients between the CRSP spread and the TAQ spread ranges from 0.8267 in 1996 to 0.9603 in 2003 for NYSE/AMEX stocks. For NASDAQ stocks, the annual average of monthly cross-sectional correlation coefficients between the CRSP spread and the TAQ spread ranges from 0.9193 to 0.9729. The cross-sectional correlations between the CRSP spread and TAQ spread are quite robust across firms with different characteristics, especially for NASDAQ firms. We also provide evidence that the simple CRSP-based spread provides a better approximation of the TAQ spread than all other low-frequency liquidity measures in cross-sectional settings. However, the CRSP spread is highly correlated with the TAQ spread in time-series settings only for NASDAQ stocks. Overall, our empirical results suggest that the simple CRSP spread could be used in lieu of the TAQ spread in academic research that focuses on cross-sectional analysis.

The rest of the paper is organized as follows. Section 2 provides the definition of each liquidity measure and presents descriptive statistics. In Section 3, we analyze the cross-sectional relation between the CRSP spread and the TAQ spread. Section 4 provides evidence on how the

[4] These two variables are available for NYSE/AMEX/NASDAQ stocks from 1993 onwards. These variables are also available from 1925 to 1942 for all NYSE/AMEX stocks and from 1982 to 1992 for most NASDAQ stocks.

CRSP spread compares with other low-frequency liquidity measures as an approximation of the TAQ spread in both cross-sectional and time-series settings. Section 5 provides a brief summary and concluding remarks.

2. Liquidity measures and descriptive statistics

In this section we introduce various liquidity measures and provide descriptive statistics.

2.1. The CRSP bid-ask spread

For NYSE/AMEX stocks, the CRSP database provides a continuous series of *Ask* (the closing ask price) and *Bid* (the closing bid price) from December 31, 1925 through February 23, 1942.[5] Between February 24, 1942 and December 27, 1992, *Ask* and *Bid* are available only in cases when a closing price is missing.[6] The CRSP database provides a continuous series of *Ask* and *Bid* for NYSE/AMEX stocks from December 28, 1992. For NASDAQ stocks, the CRSP database provides a continuous series of *Ask* and *Bid* for National Market System (NMS) securities from November 1, 1982 and all securities from June 15, 1992 with the following two exceptions: *Ask* and *Bid* are missing for 15 NMS securities in December 1982 and for all NMS securities in February 1986.[7] For NASDAQ stocks, the CRSP database provides the closing inside quotes (i.e., the highest bid and lowest ask prices). For NYSE stocks, *Ask* and *Bid* are not

[5] Jones (2002) collected the pre-1960 monthly closing bid and ask data for a subset of stocks in the Dow Jones Averages from printed sources (e.g., the *Commercial and Financial Chronicle*) and showed that spreads are cyclical and spreads predict stock returns using data for the 1900-2000 period.

[6] These items are available only when there are no trades in the *Bid or Low Price* and *Ask or High Price* fields, and only on days when there are no trades.

[7] The CRSP database provides the pre-ISSM data for NASDAQ NMS securities (the ISSM database covers only the 1987-1992 period).

the closing inside quotation. Instead, they represent the last representative quotes before the markets close for each trading day.[8]

We calculate the CRSP bid-ask spread of stock *i* on day τ using the following formula:

$$CRSP\_Spread_{i,t} = (Ask_{i,t} - Bid_{i,t})/M_{i,t}; \quad (1)$$

where $Ask_{i,t}$ is the ask price of stock *i* on day *t* from the CRSP daily data, $Bid_{i,t}$ is the bid price of stock *i* on day *t* from the CRSP daily data, and $M_{i,t}$ is the mean of $Ask_{i,t}$ and $Bid_{i,t}$. To reduce the effect of data errors and outliers, we exclude all $CRSP\_Spread_{i,t}$ that are greater than 50% of the quote midpoint. For each stock, we then calculate the monthly and yearly mean values of $CRSP\_Spread_{i,t}$ from 1993 through 2009.[9]

2.2. The TAQ bid-ask spread

We obtain intraday trade and quote data from the Trade and Quote (*TAQ*) database provided by the NYSE. We first generate national best bid and offer (NBBO) quotes and apply the following filters, which are standard in the microstructure literature (e.g., Huang and Stoll, 1996; Chung, Van Ness, and Van Ness, 2001), to clean data errors and outliers: (1) delete quotes if either the bid or ask price is negative; (2) delete quotes if either the bid or ask size is negative; (3) delete quotes if the bid-ask spread is greater than $5 or negative; (4) delete quotes if they are out of time sequence or involve an error; (5) delete before-the-open and after-the-close trades and quotes; and (6) delete quotes if they differ by more than 10% from the previous quote. Following prior studies (e.g., Bessembinder and Kaufman, 1997; Bettis, Coles, and Lemmon,

[8] In the CRSP database, *Bid* and *Ask* are set to zero if available quotes are unrepresentative of trading activity. We delete CRSP observations if both *Ask* and *Bid* are zero. These observations constitute less than 0.6% of all firm-day observations in our study sample.

[9] The vast majority of stocks (i.e., more than 96% of the stock-months) have valid CRSP bid-ask spreads on more than 15 days each month during the 1993-2009 period.

2000; Eleswarapu and Venkataraman, 2006), we calculate the effective spread of stock *i* at time τ using the following formula:

$$TAQ\_Spread_{i,\tau} = 2D_{i,\tau}(P_{i,\tau} - M_{i,\tau})/M_{i,\tau}, \quad (2)$$

where $P_{i,\tau}$ is the transaction price at time τ, $M_{i,\tau}$ is the midpoint of the most recently posted NBBO quotes (i.e., the mean of TAQ intraday NBBO ask and bid prices), and $D_{i,\tau}$ is a binary variable which equals one for buyer-initiated trades and negative one for seller-initiated trades.[10] For each stock, we first calculate the daily trade-weighted average effective spread and then the monthly and yearly mean values of the effective spread from 1993 through 2009.[11]

2.3. Other low-frequency liquidity measures

We estimate monthly and yearly values of various low-frequency liquidity measures using daily CRSP data.

2.3.1. Roll's spread

Roll (1984) shows that the bid-ask spread can be estimated by the serial covariance in price changes:

$$S = 2\sqrt{-Cov(\Delta P_t, \Delta P_{t-1})}\ , \quad (3)$$

where P_t is the closing price of a stock on day *t*. We obtain *Roll_Spread* by dividing *S* by the average price during the estimation period. Following Goyenko, Holden, and Trzcinka (2009), we assume that *Roll_Spread* = 0 when $Cov(\Delta P_t, \Delta P_{t-1}) > 0$.

[10] We estimate $D_{i,\tau}$ using the Lee and Ready (1991) algorithm modified by Bessembinder (2003b).
[11] We also find that the simple CRSP-based spread provides a better approximation of the TAQ-based *quoted* bid-ask spread than all other low-frequency liquidity measures in cross-sectional settings. For brevity, we do not tabulate the results.

2.3.2. Effective tick

Goyenko, Holden, and Trzcinka (2009) and Holden (2009) develop a proxy for the effective spread based on price clustering. Let S_t be the realization the closing effective spread on day t and assume that S_t is randomly drawn from a set of possible spreads s_j, j = 1, 2,…, J with corresponding probabilities γ_j, j = 1, 2, …, J. Let N_j be the number of trades on prices corresponding to the jth spread (j = 1, 2, …,J). The proportion of trades on prices corresponding to the jth spread is given by:

$$F_j = \frac{N_j}{\sum_{j=1}^{J} N_j} \text{ for } j = 1, 2, \ldots, J. \tag{4}$$

The unconstrained probability of the jth spread (j = 1, 2, …, J) is given by:

$$U_j = \begin{cases} 2F_j, & j = 1 \\ 2F_j - F_{j-1}, & j = 2,3,...,J-1 \\ F_j - F_{j-1}, & j = J \end{cases}. \tag{5}$$

Goyenko, Holden, and Trzcinka (2009) and Holden (2009) obtain the constrained probability of the jth spread (j = 1, 2, …, J) from:

$$\hat{\gamma}_j = \begin{cases} Min\left[Max\{U_j, 0\},\ 1\right], & j = 1 \\ Min\left[Max\{U_j, 0\},\ 1 - \sum_{k=1}^{j-1} \hat{\gamma}_k \right], & j = 2,3,...,J \end{cases}. \tag{6}$$

The effective tick is the probability-weighted average of the jth spread divided by the average price, $\overline{P}_i$, in time interval i. That is,

$$Effective\_Tick = \frac{\sum_{j=1}^{J} \hat{\gamma}_j s_j}{\overline{P}_i}. \tag{7}$$

We calculate *Eff_Tick* using daily prices from positive-volume days. Following Goyenko, Holden, and Trzcinka (2009), we also calculate a second version, *Eff_Tick2*, using daily prices from all days.

2.3.3. Gibbs estimate

Hasbrouck (2009) proposes a Gibbs estimate (labeled as *Gibbs* in our study) of the effective cost of trading that is based on daily closing prices. Hasbrouck (2009) assumes that the public information shock is normally distributed with mean of zero and variance of σ_e^2. Hasbrouck uses the Gibbs sampler to numerically estimate the Roll model parameters (i.e., *Gibbs* and σ_e^2).

2.3.4. Lesmond, Ogden, and Trzcinka (LOT) measure

Lesmond, Ogden, and Trzcinka (1999) develop an estimator of transaction cost based on the idea that non-zero returns are observed only if the inherent true return exceeds the trading cost threshold. The *LOT* model makes the following assumption regarding the relation between observed returns, R_{jt}, and unobserved "true" returns, R_{jt}^*:

$$R_{jt}^* = \beta_j R_{mt} + \varepsilon_{jt}, \tag{8}$$

Where:

$$\begin{aligned} R_{jt} &= R_{jt}^* - \alpha_{1j} \quad \text{if} \quad R_{jt}^* < \alpha_{1j} \\ R_{jt} &= 0 \quad\quad \text{if} \quad \alpha_{1j} \le R_{jt}^* \le \alpha_{2j} \\ R_{jt} &= R_{jt}^* - \alpha_{2j} \quad \text{if} \quad R_{jt}^* > \alpha_{2j}. \end{aligned} \tag{9}$$

Lesmond, Ogden, and Trzcinka (1999) estimate α_{1j} and α_{2j} using the maximum likelihood method and measure the proportional round trip transaction cost by the difference between the two estimates:

$$LOT = \alpha_{2j} - \alpha_{1j}. \tag{10}$$

2.3.5. Zeros

Lesmond, Ogden, and Trzcinka (1999) suggest that the proportion of days with zero returns can be a proxy for transaction costs for two reasons: (1) stocks with higher trading costs are more likely to have zero-volume days and thus zero-return days and (2) stocks with higher trading costs are also more likely to have zero returns even in positive-volume days. As in Lesmond, Ogden, and Trzcinka (1999), we employ two definitions of the proportion of days with zero returns: *Zeros* = (Number of days with zero returns)/(Number of trading days) and *Zeros2* = (Number of positive-volume days with zero returns)/(Number of positive-volume trading days).

2.3.6. Amihud illiquidity measure

Amihud (2002) suggests a measure of trading cost (i.e., illiquidity) that is based on the price impact of trading:

$$Illiquidity = 1{,}000 \cdot Average\left(\frac{|r_t|}{Volume_t} \right), \tag{11}$$

where r_t and $Volume_t$ denote stock return and dollar trading volume, respectively, on day *t*.

2.3.7. Pastor and Stambaugh (PS) measure

Pastor and Stambaugh (2003) employ the following regression model to measure the liquidity of stock *i*, $Gamma_i$:

$$r^{e}_{i,d+1} = \theta_i + \phi_i r_{i,d} + Gamma_i \cdot sign(r^{e}_{i,d}) \cdot v_{i,d} + \varepsilon_{i,d+1} \quad (12)$$

where $r_{i,d}$ is stock *i*'s return on day *d*, $r^{e}_{i,d} = r_{i,d} - r_{m,d}$ where $r_{m,d}$ is the CRSP value-weighted market return on day *d*, and $v_{i,d}$ is stock *i*'s dollar trading volume on day *d*. If a stock were not liquid, its order flow would be accompanied by a return that is likely to be reversed. Hence the greater the expected reversal for a given dollar trading volume, the lower the stock's liquidity. *Gamma* measures the return reversal due to the previous day's order flow shock.

2.4. Descriptive statistics

Our study sample includes all NYSE, AMEX, and NASDAQ stocks with available CRSP and TAQ data over the period from 1993 to 2009. We calculate the monthly and yearly values of each liquidity measure that results in 1,471,945 firm-month observations and 137,751 firm-year observations.[12] In Table 1, Panel A and Panel B present descriptive statistics on *TAQ_Spread*, *CRSP_Spread*, and other low-frequency liquidity measures that are based on their monthly and yearly observations, respectively. Panel A shows that the mean value (0.0293) of *CRSP_Spread* is greater than the mean value (0.0203) of *TAQ_Spread*. The mean value (0.0288) of *Roll_Spread* is comparable to that of the *CRSP_Spread*, but greater than that of the *TAQ_Spread*. The mean values of *Eff_Tick* and *Eff_Tick2* are 0.016 and 0.0151, respectively. The mean values of *Gibbs* and *LOT* are 0.0135 and 0.0434, respectively. The mean values of *Zeros* and *Zeros2* are 0.1263 and 0.1179, which are quite similar. Panel B shows that descriptive statistics from yearly data are similar to those from monthly data.[13]

[12] We exclude firm-month observations with less than 10 trading days in any month and firm-year observations with less than 40 trading days in any year. We obtain qualitatively similar results when we apply different filters.

[13] We winsorize the monthly/yearly observations of low-frequency liquidity measures at the 1st and 99th percentiles to

3. Spreads: CRSP versus TAQ

In this section, we examine the cross-sectional correlation between the CRSP spread and the TAQ spread.

3.1. Results for the whole sample

We calculate the monthly mean values of CRSP spreads and TAQ spreads, respectively, for each stock and obtain their cross-sectional mean and median values in each year. We also calculate the cross-sectional correlation coefficient (ρ) between the mean CRSP and TAQ spreads in each month and obtain its mean value in each year. Panel A and Panel B (see the left half of each panel) in Table 2 show the results for NYSE/AMEX and NASDAQ stocks, respectively. To assess whether our results are sensitive to data aggregation methods, we replicate the above results using yearly data. That is, we calculate yearly mean values of CRSP spreads and TAQ spreads for each stock and obtain their cross-sectional mean and median values in each year. We also calculate the cross-sectional correlation coefficient between the mean CRSP and TAQ spreads in each year. The right half of Panels A and B shows these results. The last two rows show the *t*- and *z*-statistics for testing whether the monthly/yearly mean and median values of CRSP spreads are significantly greater than those of TAQ spreads over the 17-year study period.

Panel A of Table 2 shows that for NYSE/AMEX stocks, the mean value of CRSP spreads is larger than the mean value of TAQ spreads in most years (except from 2005 to 2007) and their mean difference during the 17-year period is statistically significant, regardless of whether we use the monthly or yearly data aggregation. The median value of CRSP spreads is greater than

reduce the influence of outliers and possible data errors.

the median value of TAQ spreads in all years and their difference during the 17-year period is statistically significant.

We find similar results for NASDAQ stocks. The mean value of CRSP spreads is larger than the mean value of TAQ spreads in most years (except from 2003 to either 2006 or 2007, depending on the aggregation method) and their mean difference during the 17-year period is statistically significant. The median value of CRSP spreads is larger than the median value of TAQ spreads in most years (except from 2002 to 2006) and their difference based on monthly data during the 17-year period is statistically significant.[14]

Although the mean/median values of CRSP spreads are different from those of TAQ spreads for both NYSE/AMEX and NASDAQ stocks, they are highly correlated with each other. For NYSE/AMEX stocks, Panel A of Table 2 shows that the correlation coefficients between TAQ spreads and CRSP spreads are all higher than 0.82 (0.83) for monthly (yearly) data. Panel B shows that the correlation coefficients between TAQ spreads and CRSP spreads are even higher for NASDAQ stocks (i.e., at least 0.91 for monthly data and 0.93 for yearly data). The higher correlation between TAQ spreads and CRSP spreads for NASDAQ stocks may largely be attributed to the fact that the CRSP database reports the closing inside quotes of NASDAQ stocks, whereas it reports the last representative quotes of NYSE/AMEX stocks.

[14] The smaller CRSP spread (relative to the TAQ effective spread) during 2002-2007 is puzzling. To shed further light on this issue, we cluster NASDAQ stocks into quintiles based on market capitalization during 2003-2006 and compare the CRSP and TAQ spreads for stocks within each quintile. The results show that the CRSP spread is smaller than the TAQ spread across all size quintiles, indicating that the smaller CRSP spread is not due to some size-related reasons. Note that the CRSP spread is the quoted spread at market close, whereas the TAQ spread is the trade-weighted effective spread calculated from all relevant intraday inside quotes. There are at least two possible reasons why the quoted spread at market close might have decreased around 2003. The decrease in closing spreads might have resulted from the regulatory enforcement against closing price manipulation (Commerton-Forde and Putnins, 2011). Pagano and Schwartz (2003) provide evidence that the introduction of closing call auction on the Euronext Paris improves market quality at market-close and during the pre-close trading intervals. In addition, Archipelago's introduction of closing auctions in January 2003 and the introduction of an electronic closing call auction on NASDAQ in March 2004 have exerted the similar impact on closing spreads (Pagano, Peng, and Schwartz, 2013).

Consistent with prior research, we find that the TAQ spread generally declined during our study period.[15] For NASDAQ stocks (see Panel B in Table 2), the trend in the CRSP spread is similar to the trend in the TAQ spread. For NYSE/AMEX stocks (see Panel A in Table 2), there is less similarity in the time-series trend between the CRSP spread and the TAQ spread. For example, the TAQ spread decreases from 1993 to 1997, whereas the CRSP spread increases during the same period.

To shed additional light on the relation between the CRSP spread and the TAQ spread, we cluster our sample stocks into five portfolios according to the market value of equity at the end of each month and year (i.e., the number of shares outstanding times share price at the end of each month and year). We then aggregate all firm-month/firm-year observations within each portfolio across months/years. Panel C in Table 2 shows the mean and median values of the CRSP spread and the TAQ spread for each market capitalization quintile. Consistent with the finding of prior studies (e.g., Chung, McInish, Wood, and Wyhowski, 1995), we find that spreads decrease monotonically with market capitalizations. The CRSP spreads are greater than the TAQ spreads in all quintiles, except that the median value of the CRSP spread is very close to the median value of the TAQ spread in quintile 5 of NASDAQ stocks.

Panel C of Table 2 also shows that for NASDAQ stocks, the correlation coefficients between the CRSP spread and the TAQ spread are consistently greater than 0.86 across all quintiles, indicating that the CRSP spread is highly correlated with the TAQ spread across firms with different market capitalizations. For NYSE/AMEX stocks, the correlation coefficients between the CRSP spread and the TAQ spread range from 0.65 to 0.88 across all quintiles.

[15] Larger percentage spreads in 2008 and 2009 relative to those in 2007 could be attributed, at least in part, to lower share prices during and after the subprime crisis.

Finally, we note that the correlation coefficients between the CRSP and TAQ spreads are relatively higher for firms with smaller market capitalizations.

4. Comparison with other low-frequency liquidity measures

In this section we analyze how the CRSP spread compares with other low-frequency liquidity measures as an approximation of the TAQ spread.

4.1. Cross-sectional correlation

We calculate the monthly cross-sectional correlation coefficients between *TAQ_Spread* and each low-frequency liquidity measure for NYSE/AMEX stocks and NASDAQ stocks, respectively,[16] and obtain their mean values for the entire study period 1993-2009 as well as for three sub-periods 1993-1996, 1997-2000, and 2001-2009.[17] We show the results in Table 3. The results indicate that all the mean correlation coefficients are statistically significant at the 5% level, except the mean correlation coefficient between *PS* and *TAQ_Spread* for NYSE/AMEX stocks in the 2001-2009 period.

To compare the mean correlation coefficients between *TAQ_Spread* and each of the low-frequency liquidity measures during a given time period, we conduct the *t*-test using the time-series observations of the monthly cross-sectional correlation coefficients.[18] We put *, **, or *** on the correlation coefficient if it is significantly greater than all other correlation coefficients at the 10%, 5%, or 1% level. If such correlation coefficient does not exist, we put † on the highest

[16] We find qualitatively similar but weaker results (e.g., smaller *t*-values) when we conduct tests using yearly observations.
[17] These sub-periods correspond to the tick-size regimes of $1/8, $1/16, and $0.01, respectively.
[18] We adjust standard errors for autocorrelation with a Newey-West correction using four lags in the *t*-tests in Table 3 and Table 4.

correlation coefficient and ‡ on the correlation coefficients that are not significantly different from the highest correlation coefficient at the 5% level according to the *t*-test.

The results in Table 3 show that for the entire sample period, the mean correlation coefficient between *TAQ_Spread* and *CRSP_Spread* is 0.9020 for NYSE/AMEX stocks and 0.9515 for NASDAQ stocks, which are significantly greater than the mean correlation coefficients between *TAQ_Spread* and other low-frequency liquidity measures. We find similar results for sub-periods with one exception: the correlation coefficient (0.8805) between *CRSP_Spread* and the *TAQ_Spread* for NYSE/AMEX stocks is greater than but not significantly different from the correlation coefficient (0.8790) between the *Gibbs* and *TAQ_Spread* during the 1993-1996 period. Overall, these results indicate that *CRSP_Spread* provides a better approximation of the TAQ-based spread than any other low-frequency liquidity measures suggested in the literature, especially for NASDAQ stocks.

4.2. Cross-sectional correlation: Results from subsample analysis

We conduct subsample analysis to examine whether the cross-sectional correlation between the CRSP spread and the TAQ spread varies across firms with different market capitalizations, return volatilities, and *TAQ_Spread* and also to provide further evidence regarding whether the CRSP spread provides a better approximation of the TAQ spread than other low-frequency liquidity measures. We measure a firm's market capitalization by the product of the number of shares outstanding and share price at the end of each month and return volatility by the standard deviation of daily stock returns in each month.

In each month, we sort firms according to market capitalization, return volatility, and *TAQ_Spread*, respectively, and group them into five portfolios (quintiles). We first calculate the

cross-sectional correlation coefficient between the CRSP spread and the TAQ spread for stocks within each quintile in each month. We then calculate the mean value of the correlation coefficients within each quintile across months during the entire study period of 1993 to 2009. Panel A, Panel B, and Panel C in Table 4 show the results for market capitalizations, return volatility, and *TAQ_Spread* quintiles, respectively. In each panel, we report the results for NYSE/AMEX stocks and NASDAQ stocks separately.

In Panels A and B in Table 4, the results for NASDAQ stocks show that the correlation coefficient between the CRSP spread and the TAQ spread is consistently greater than 0.86 across all quintiles, regardless of whether we sort firms based on market capitalizations (in Panel A) or return volatility (in Panel B). For NYSE/AMEX stocks, the correlation coefficient between the CRSP spread and the TAQ spread varies significantly across quintiles from 0.65 to 0.90. The correlation coefficients between the CRSP spread and the TAQ spread are relatively higher for firms with smaller market capitalizations and higher return volatility.

We also test for each quintile whether the time-series mean values of the monthly cross-sectional correlations between *TAQ_Spread* and the CRSP spread are significantly greater than the corresponding figures between *TAQ_Spread* and other low-frequency liquidity measures. The results indicate that the correlation coefficient between *CRSP_Spread* and *TAQ_Spread* is significantly greater than the correlation coefficients between *TAQ_Spread* and all other low-frequency liquidity measures at the 1% significance level across all market-capitalization quintiles (Panel A in Table 4) and return volatility quintiles (Panel B in Table 4) for both NYSE/AMEX and NASDAQ stocks. These results provide further evidence that the simple CRSP spread provides a better approximation of the TAQ-based spread than any other low-frequency liquidity measures suggested in the literature.

Panel C of Table 4 shows the mean correlation coefficient between *TAQ_Spread* and each low-frequency liquidity measure for stocks within each *TAQ-Spread* quintile. Compared to the results in Panel A and Panel B, the correlation coefficients are much smaller. The correlation coefficients between the TAQ spread and the CRSP spread range from 0.49 to 0.86 for NASDAQ stocks and from 0.21 to 0.83 for NYSE/AMEX stocks.[19] Similarly, the correlation coefficients between the TAQ spread and other low-frequency liquidity measures are much smaller than the corresponding values in Panel A and Panel B.

In Panel C of Table 4, the *t*-test results indicate that *CRSP_Spread* is more strongly related to *TAQ_Spread* than all other low-frequency liquidity measures across the TAQ spread quintiles for both NYSE/AMEX and NASDAQ stocks, except quintile 1 for NYSE/AMEX stocks. For NYSE/AMEX stocks in quintile 1, although *Eff_Tick* has the highest correlation (0.4310) with *TAQ_Spread*, the correlation coefficient (0.4139) between *CRSP_Spread* and the *TAQ_Spread* and the correlation coefficient (0.4216) between *Eff_Tick2* and the *TAQ_Spread* are not significantly different from the correlation coefficient between *Eff_Tick* and *TAQ_Spread*. Overall, the results in Table 4 suggest that the CRSP spread provides a better approximation of the TAQ spread than all other low-frequency liquidity measures across firms with different characteristics, especially for NASDAQ stocks.

4.3. Time-series correlation

To further assess how *CRSP_Spread* compares with other low-frequency liquidity measures as an approximation of the TAQ spread, we estimate the time-series correlation

[19] When quintiles are formed based on the TAQ spread, the cross-sectional variation in the TAQ spread within each quintile is relatively smaller than the corresponding variation in the TAQ spread when quintiles are formed based on other firm attributes (e.g., market capitalization or return volatility). As a consequence, the correlations between the TAQ spread and low-frequency liquidity measures in Panel C of Table 4 are much lower than those in Panel A and Panel B.

between the TAQ spread and each low-frequency liquidity measure. We use a method employed by Goyenko, Holden, and Trzcinka (2009) to calculate the time-series correlation coefficients. Specifically, we first calculate the monthly cross-sectional mean values of each liquidity measure for NYSE/AMEX stocks and NASDAQ stocks, respectively. Then we calculate the time-series correlation coefficient between the monthly mean TAQ spreads and the monthly mean values of each low-frequency liquidity measure for our entire study period 1993-2009, as well as for the three sub-periods 1993-1996, 1997-2000, and 2001-2009.

Table 5 shows that the time-series correlation coefficient between the CRSP spread and the TAQ spread is 0.9738 for the entire sample period, and 0.97, 0.9126, and 0.9342 for the three sub-periods for NASDAQ stocks. The results also show that the time-series correlation coefficients between the TAQ spread and *Roll_Spread*, *Eff_Tick*, *Eff_Tick2*, and *Zeros* are all greater than 0.9 for the entire sample period, 1993-2009. For NASDAQ stocks, the Fisher's *Z*-test results show that the time-series correlation between *TAQ_Spread* and the CRSP spread is significantly greater than that between *TAQ_Spread* and any other low-frequency liquidity measures for the entire study period. [20] We find generally similar results for each sub-period, except the 2001-2009 period.[21]

For NYSE/AMEX stocks, the correlation coefficient (0.7906) between the CRSP spread and the TAQ spread for the entire study period is lower than the correlation coefficients (0.8411 and 0.8243) between the TAQ spread and two other low-frequency liquidity measures (i.e., *Eff_Tick* and *Eff_Tick2*). The results of Fisher's *Z*-test show however that the correlation

[20] We conduct Fisher's *Z*-test to compare the time-series correlation coefficients between *TAQ_Spread* and each of the low-frequency liquidity measures during a given time period. We put *, **, or *** on the correlation coefficient if it is significantly greater than all other correlation coefficients at the 10%, 5%, or 1% level. If such correlation coefficient does not exist, we put † on the highest correlation coefficient and ‡ on the correlation coefficients that are not significantly different from the highest correlation coefficient at the 5% level according to Fisher's *Z*-test.

[21] During this sub-period, the correlation coefficient between the TAQ spread and the CRSP spread is smaller than both the correlation coefficient between the TAQ spread and *Eff_Tick* and the correlation coefficient between the TAQ spread and *Eff_Tick2*, although the differences are not statistically significant.

coefficient between *Eff_Tick* and *TAQ_Spread* is not significantly different from the correlation coefficients between *TAQ_Spread* and *CRSP_Spread/Eff_Tick2* at the 5% level.

From the sub-period results for NYSE/AMEX stocks, we find that *Eff_Tick*, *Eff_Tick2*, *Gibbs*, and *Illquidity* provide better approximations of *TAQ_Spread* than other liquidity measures during the 1993-1996 period; *CRSP_Spread*, *Roll_Spread*, and *Gibbs* provide better approximations of *TAQ_Spread* than other measures during the 1997-2000 period; and *CRSP_Spread*, *Eff_Tick*, and *Eff_Tick2* provide better approximations of *TAQ_Spread* than other measures during the 2001-2009 period. We note that the correlation coefficient between *CRSP_Spread* and *TAQ_Spread* during the 1993-1996 period is negative, which is consistent with the result in Table 2 that the TAQ spread decreased and the CRSP spread increased during the same period.[22] We also note that *CRSP_Spread* remains as one of the best approximations of *TAQ_Spread* in more recent periods (i.e., from 1997 onwards) even for the NYSE/AMEX stocks in time-series setting.

Overall, the results suggest that for NASDAQ stocks, the CRSP spread provides a better approximation of the TAQ spread than most other low-frequency liquidity measures in the time-series setting. However, for NYSE/AMEX stocks, the CRSP spread generally provides a poorer approximation of the TAQ spread than for NASDAQ stocks, and some low-frequency liquidity measures (e.g., *Eff_Tick* and *Eff_Ticks*) provide better approximations of the TAQ spread than the CRSP spread.

[22] It is unclear why the CRSP spread and the TAQ spread are negatively related for NYSE/AMEX stocks during 1993-1996. This anomalous result may be attributed, at least in part, to the fact that the CRSP provides only the last representative quotes for NYSE/AMEX stocks, whereas it provides the closing inside quotes for NASDAQ stocks. To shed further light on this issue, we cluster NYSE/AMEX stocks into quintiles based on market capitalization, share price, return volatility, and *TAQ_Spread* during 1993-1996 and examine the time-series pattern of the CRSP and TAQ spreads for stocks within each quintile. The results show that the CRSP spread increased during 1993-1996 across all market capitalization/share price/return volatility/*TAQ_Spread* quintiles, indicating that the negative relation is not limited to only certain stocks.

4.4. Time-series correlation: Results from subsample analysis

Panels A, B, and C in Table 6 show the results of the time-series correlation analysis for subsamples formed by market capitalizations, return volatility, and the TAQ spread, respectively, for the entire study period 1993-2009. Panel A shows that for NASDAQ stocks, the correlation coefficients between the CRSP spread and the TAQ spread are all greater than 0.95 across all firm size quintiles. Some other low-frequency liquidity measures are also highly correlated with the TAQ spread. For example, *Eff_Tick* and *Eff_Tick2* are highly correlated with the TAQ spread in the subsample of large firms, while *Roll_Spread* and *Gibbs* are highly correlated with the TAQ spread in the subsample of small firms. However, only *CRSP_Spread* remains as a "winner or co-winner" across all five subsamples according to Fisher's *Z*-test, which is consistent with the finding in Table 5 that the CRSP spread provides a better approximation of the TAQ spread than most other low-frequency liquidity measures in the time-series setting for NASDAQ stocks.

For NYSE/AMEX stocks, *Eff_Tick* and *Eff_Tick2* have higher time-series correlations with the TAQ spread than the CRSP spread and other low-frequency liquidity measures in most firm size quintiles. Specifically, the Fisher's *Z*-test results show that *Eff_Tick* and *Eff_Tick2* are the "co-winners" in four of the five subsamples, which is consistent with the finding in Table 5 that *Eff_Tick* and *Eff_Ticks* provide better approximations of the TAQ spread than other liquidity measures in the time-series setting for NYSE/AMEX stocks.

Panel B of Table 6 shows that for NASDAQ stocks, the correlation coefficients between the TAQ spread and the CRSP spread are all above 0.969 across the volatility quintiles and significantly greater than the correlation coefficients between the TAQ spread and all other low-frequency liquidity measures across all return volatility quintiles according to Fisher's *Z*-test. Some other low-frequency liquidity measures also have high correlations with the TAQ spread.

For example, *Eff_Tick* and *Eff_Tick2* are highly correlated with the TAQ spread across all volatility subsamples, while *Roll_Spread* is highly correlated with the TAQ spread in the subsample of high-volatility firms.

For NYSE/AMEX stocks, Panel B of Table 6 shows that *Eff_Tick* and *Eff_Tick2* have higher time-series correlations with the TAQ spread than the CRSP spread and other low-frequency liquidity measures in most return volatility quintiles. Specifically, the Fisher's *Z*-test results show that *Eff_Tick* and *Eff_Tick2* are the "co-winners" in four of the five subsamples, which is consistent with our previous finding that *Eff_Tick* and *Eff_Ticks* provide good approximations of the TAQ spread in the time-series setting for NYSE/AMEX stocks. For NYSE/AMEX stocks, Panel B also shows that *CRSP_Spread* remains as a "co-winner" in larger volatility quintiles (three of the five subsamples).

Panel C of Table 6 shows that for NASDAQ stocks, the time-series correlations between the CRSP spread and the TAQ spread are above 0.96 across all the quintiles formed by the TAQ spread. Some other low-frequency liquidity measures also have high (e.g., above 0.90) correlations with the TAQ spread. For example, *Eff_Tick* and *Eff_Tick2* are highly correlated with the TAQ spread in the subsample of firms with small TAQ spreads, while *Roll_Spread* and *Gibbs* are highly correlated with the TAQ spread in the subsample of firms with large TAQ spreads. However, only *CRSP_Spread* remains as a "winner or co-winner" across all five subsamples according to Fisher's *Z*-test, which is consistent with our previous finding that the CRSP spread provides a better approximation of the TAQ spread than most other low-frequency liquidity measures in the time-series setting for NASDAQ stocks.

For NYSE/AMEX stocks, Panel C shows that *Eff_Tick* and *Eff_Tick2* have higher time-series correlations with the TAQ spread than the CRSP spread and other low-frequency liquidity

measures in most *TAQ_Spread* quintiles. Specifically, the Fisher's *Z*-test results show that *Eff_Tick* and *Eff_Tick2* are the "co-winners" in four of the five subsamples, which is consistent with our previous finding.

Taken together, our subsample analysis in Panels A to C in Table 6 suggests that the CRSP spread provides a better approximation of the TAQ spread than other low-frequency liquidity measures for most NASDAQ stocks, while other low-frequency liquidity measures (i.e., *Eff_Tick* and *Eff_Tick2*) provide a better approximation of the TAQ spread than the CRSP spread for most NYSE/AMEX stocks.

5. Summary and concluding remarks

The bid-ask spread has been widely used as a measure of stock market liquidity and information asymmetry in prior research. In this study we show the usefulness of a simple bid-ask spread measure that can be calculated using only daily data provided by the CRSP. We show that this simple measure provides a good approximation of the bid-ask spread from intraday data. The CRSP-based spread is highly correlated with the TAQ-based effective spread in cross-sectional settings, especially for NASDAQ stocks. We also provide evidence that the simple CRSP-based spread provides a better approximation of the TAQ-based spread than all other low-frequency liquidity measures in cross-sectional settings. We also show that the CRSP-based spread is highly correlated with the TAQ-based spread in time-series settings for NASDAQ stocks. However, the CRSP-based spread provides a poorer approximation of the TAQ-based spread for NYSE/AMEX stocks than for NASDAQ stocks. The lower correlations between the CRSP spread and the TAQ spread for NYSE/AMEX stocks may be attributed in part to that the

CRSP database provides only the last representative bid and ask quotes for NYSE/AMEX stocks whereas it provides the closing *inside* bid and ask quotes for NASDAQ stocks.

Our low-frequency liquidity measure does not require a sophisticated estimation method like some other low-frequency liquidity measures, making it easier for both researchers and practitioners to use. The CRSP-based spread would be particularly useful to those who want to incorporate stock market liquidity in their research without having to go through the process that is required for the TAQ database.

References

Ali, A., Chen, T., Radhakrishnan, S., 2007. Corporate disclosures by family firms. Journal of Accounting and Economics 44, 238–286.

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

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid-ask spread. Journal of Financial Economics 17, 223–249.

Banerjee, S., Gatchev, V., Spindt, P. A., 2007. Stock market liquidity and firm dividend policy. Journal of Financial and Quantitative Analysis 42, 369–397.

Barclay, M. J., Christie, W. G., Harris, J. H., Kandel, E., Schultz, P. H., 1999. Effects of market reform on the trading costs and depths of NASDAQ stocks. Journal of Finance 54, 1–34.

Bessembinder, H., 2003a. Trade execution costs and market quality after decimalization. Journal of Financial and Quantitative Analysis 38, 747–778.

Bessembinder, H., 2003b. Issues in assessing trade execution costs. Journal of Financial Markets 6, 233–257.

Bessembinder, H., Kaufman, H., 1997. A comparison of trade execution costs for NYSE and NASDAQ listed stocks. Journal of Financial and Quantitative Analysis 32, 287–310.

Bettis, C., Coles, J., Lemmon, M., 2000. Corporate policies restricting trading by insiders. Journal of Financial Economics 57, 191–220.

Brockman, P., Chung, D. Y., Yan, X., 2009. Block ownership, trading activity, and market liquidity. Journal of Financial and Quantitative Analysis 44, 1403–1426.

Brockman, P., Howe, J. S., Mortal, S., 2008. Stock market liquidity and the decision to repurchase. Journal of Corporate Finance 14, 446–459.

Bushee, B. J., Core, J. E., Guay, W., Hamm, S., 2010. The role of the business press as an information intermediary. Journal of Accounting Research 48, 1–19.

Chang, H., Chen, J., Liao, W. M., Mishra, B. K., 2006. CEOs'/CFOs' swearing by the numbers: Does it impact share price of the firm. The Accounting Review 81, 1–27.

Cheng, M., Dhaliwal, D. S., Neamtiu, M., 2011. Asset securitization, securitization recourse, and information uncertainty. The Accounting Review 86, 541–568.

Christie, W. G., Schultz, P. H., 1994. Why do NASDAQ market makers avoid odd-eighth quotes? Journal of Finance 49, 1813–1840.

Chung, K. H., Elder, J., Kim, J., 2010. Corporate governance and liquidity. Journal of Financial and Quantitative Analysis 45, 265–291.

Chung, K. H., McInish, T., Wood, R., Wyhowski, D., 1995. Production of information, information asymmetry, and the bid-ask spread: Empirical evidence from analysts' forecasts. Journal of Banking and Finance 19, 1025–1046.

Chung, K. H., Van Ness, R. F., Van Ness, R. A., 2001. Can the treatment of limit orders reconcile the differences in trading costs between NYSE and NASDAQ issues? Journal of Financial and Quantitative Analysis 36, 267–286.

Chung, K. H., Zhang, H., 2011. Corporate governance and institutional ownership. Journal of Financial and Quantitative Analysis 46, 247–273.

Coller, M., Yohn, T., 1997. Management forecasts and information asymmetry: An examination of bid-ask spreads. Journal of Accounting Research 35, 181–191.

Comerton-Forde, C., Putnin, T. J., 2011. Measuring closing price manipulation. Journal of Financial Intermediation 20, 135–158.

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

Eckbo, B. E., Norli, O., 2002. Pervasive liquidity risk. Working paper, Dartmouth College. Available at <http://ssrn.com/abstract=996069>.

Eleswarapu, V., Venkataraman, K., 2006. The impact of legal and political institutions on equity trading costs: A cross–country analysis. Review of Financial Studies 19, 1081–1111.

Falkenstein, E., 1996. Preferences for stock characteristics as revealed by mutual fund portfolio holdings. Journal of Finance 51, 111–135.

Fang, V. W., Noe, T., Tice, S., 2009. Stock market liquidity and firm value. Journal of Financial Economics 94, 150–169.

Fang, V. W., Tian, X., Tice, S., 2013. Does stock liquidity enhance or impede firm innovation? Working paper, University of Minnesota. Available at <http://ssrn.com/abstract= 1746399>.

Frieder, L., Martell, R., 2006. On capital structure and the liquidity of a firm's stock. Working paper, Purdue University. Available at <http://ssrn.com/abstract= 880421>.

Greenstein, M., Sami, H., 1994. The impact of the SEC's segment disclosure requirement on bid-ask spreads. The Accounting Review 69, 179–199.

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

Guo, R., Lev, B., Zhou, N., 2004. Competitive costs of disclosure by biotech IPOs. Journal of Accounting Research 42, 319–355.

Hasbrouck, J., 2009. Trading costs and returns for US equities: estimating effective costs from daily data. Journal of Finance 64, 1445–1477.

Healy, P., Hutton, A., Palepu, K., 1999. Stock performance and intermediation changes

surrounding sustained increases in disclosure. Contemporary Accounting Research 16, 485–520.

Heflin, F., Shaw, K. W., 2000. Blockholder ownership and market liquidity. Journal of Financial and Quantitative Analysis 35, 621–633.

Holden, C. W., 2009. New low-frequency spread measures. Journal of Financial Markets 12, 778–813.

Huang, J., 2013. Dynamic liquidity preferences of mutual funds. Working paper, National University of Singapore. Available at <http://ssrn.com/abstract= 967553 >.

Huang, R. D., Stoll, H. R., 1996. Dealer versus auction markets. Journal of Financial Economics 41, 313–357.

Jayaraman, S., 2008. Earnings volatility, cash flow volatility, and informed trading. Journal of Accounting and Economics 46, 809–851.

Jayaraman, S., Milbourn, T., 2012. The role of stock liquidity in executive compensation. The Accounting Review 87, 537–563.

Jones, C. M., 2002. A century of stock market liquidity and trading costs. Working paper, Columbia University.

Kim, O., Verrecchia, R. E., 1994. Market liquidity and volume around earning announcements. Journal of Accounting and Economics 17, 41–67.

Lee, C., Ready, M. J., 1991. Inferring trade direction from intraday data. Journal of Finance 46, 733–746.

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

Leuz, C., 2003. IAS versus U.S. GAAP: Information asymmetry–based evidence from Germany's

new market. Journal of Accounting Research 41, 445–472.

Leuz, C., Verrecchia, R. E., 2000. The economic consequence of increased disclosure. Journal of Accounting Research 38, 91–124.

Lipson, M. L., Mortal, S., 2009. Capital structure decisions and equity market liquidity. Journal of Financial Markets 12, 611–644.

Mohd, E., 2005. Accounting for software development costs and information asymmetry. The Accounting Review 80, 1211–1231.

Pagano, M. S., Schwartz, R. A., 2003, A closing call's impact on market quality at Euronext Paris, Journal of Financial Economics 68, 439–484.

Pagano, M. S., Peng, L., Schwartz, R. A., 2013, A call auction's impact on price formation and order routing: Evidence from the NASDAQ stock market, Journal of Financial Markets, forthcoming.

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

Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficiency market. Journal of Finance 39, 1127–1139.

Sarin, A., Shastri, K. A., Shastri, K., 2000. Ownership structure and stock market liquidity. Working paper, University of Pittsburgh.

Spiegel, M. I., Wang, X., 2005. Cross-sectional variation in stock returns: Liquidity and idiosyncratic risk. Working paper, Yale University. Available at <http://ssrn.com/abstract= 709781>.

Welker, M., 1995. Disclosure policy, information asymmetry, and liquidity in equity markets. Contemporary Accounting Research 11, 801–828.

Table 1
Descriptive statistics

Panel A and Panel B present descriptive statistics on *TAQ_Spread*, *CRSP_Spread*, and other low-frequency liquidity measures that are based on their monthly and yearly observations, respectively. We calculate the CRSP bid-ask spread of stock *i* on day τ using the following formula: $CRSP\_Spread_{i,t} = (Ask_{i,t} - Bid_{i,t})/M_{i,t}$; where $Ask_{i,t}$ is the ask price of stock *i* on day *t*, $Bid_{i,t}$ is the bid price of stock *i* on day *t*, and $M_{i,t}$ is the mean of $Ask_{i,t}$ and $Bid_{i,t}$. We calculate the effective spread of stock *i* at time τ using the following formula: $TAQ\_Spread_{i,\tau} = 2D_{i,\tau}(P_{i,\tau} - M_{i,\tau})/M_{i,\tau}$; where $P_{i,\tau}$ is the transaction price, $M_{i,\tau}$ is the midpoint of the most recently posted NBBO quotes, and $D_{i,\tau}$ is a binary variable which equals one for buyer-initiated trades and negative one for seller-initiated trades. *Roll_Spread* is calculated based on Roll (1984). *Eff_Tick* and *Eff_Tick2* are constructed based on Goyenko, Holden and Trzcinka (2009) and Holden (2009). *Gibbs* is estimated based on Hasbrouck (2009). *LOT*, *Zeros*, and *Zeros2* are constructed based on Lesmond, Ogden, and Trzcinka (1999). *Illiquidity* is the illiquidity measure developed by Amihud (2002). *PS* is the Pastor and Stambaugh (2003) gamma measure. We calculate the monthly and yearly values of each liquidity measure that results in 1,471,945 firm-month observations and 137,751 firm-year observations.

	TAQ_ Spread	*CRSP_ Spread*	*Roll_ Spread*	*Eff_ Tick*	*Eff_ Tick2*	*Gibbs*	*LOT*	*Zeros*	*Zeros2*	*Illiquidity*	*PS*
Panel A. Monthly data											
Mean	0.0203	0.0293	0.0288	0.0160	0.0151	0.0135	0.0434	0.1263	0.1179	0.0031	-0.0000
Standard deviation	0.0273	0.0389	0.0576	0.0540	0.0477	0.0130	0.0362	0.1402	0.1327	0.0164	0.0007
1st percentile	0.0004	0.0005	0.0000	0.0000	0.0000	0.0013	0.0056	0.0000	0.0000	0.0000	-0.0021
Median	0.0101	0.0158	0.0116	0.0033	0.0033	0.0093	0.0327	0.0909	0.0769	0.0001	-0.0000
99th percentile	0.1395	0.1980	0.3001	0.2164	0.1957	0.0682	0.1885	0.5714	0.5385	0.0628	0.0019
Panel B. Yearly data											
Mean	0.0226	0.0328	0.0272	0.0158	0.0143	0.0123	0.0502	0.1335	0.1227	0.0037	0.0000
Standard deviation	0.0302	0.0449	0.0545	0.0379	0.0316	0.0172	0.0383	0.1237	0.1105	0.0159	0.0001
1st percentile	0.0005	0.0006	0.0000	0.0000	0.0000	0.0008	0.0076	0.0000	0.0000	0.0000	-0.0003
Median	0.0114	0.0179	0.0095	0.0041	0.0039	0.0062	0.0393	0.0965	0.0918	0.0001	0.0000
99th percentile	0.1477	0.2219	0.3170	0.2775	0.2251	0.0858	0.1990	0.5005	0.4437	0.0729	0.0006

Table 2
Correlation between the CRSP spread and the TAQ spread

We calculate the monthly mean values of CRSP spreads and TAQ spreads, respectively, for each stock and obtain their cross-sectional mean and median values in each year. We also calculate the cross-sectional correlation coefficient (ρ) between the mean CRSP and TAQ spreads in each month and obtain its mean value in each year. Panel A and Panel B (see the left half of each panel) show the results for our NYSE/AMEX and NASDAQ stocks, respectively. To assess whether our results are sensitive to data aggregation methods, we replicate the above results using yearly data. That is, we calculate yearly mean values of CRSP spreads and TAQ spreads for each stock and obtain their cross-sectional mean and median values in each year. We also calculate the cross-sectional correlation coefficient between the mean CRSP and TAQ spreads in each year. The right half of Panel A and Panel B shows these results. The last two rows show the *t*- and *z*-statistics for testing the equality of mean and median values between the CRSP and TAQ spreads over the 17-year study period. We also cluster our sample stocks into five portfolios according to the market value of equity at the end of each month and year (i.e., the number of shares outstanding times share price at the end of each month and year). We then aggregate all firm-month/firm-year observations within each portfolio across months/years. Panel C shows the mean and median values of the CRSP spread and the TAQ spread for each market capitalization quintile and the correlation coefficients between the CRSP spread and the TAQ spread. *** and * denote statistical significance at the 1% and 10% level.

Panel A: NYSE/AMEX stocks

		Monthly data				Yearly data			
Year		*CRSP_ Spread*	*TAQ_ Spread*	ρ	Number of obs.	*CRSP_ Spread*	*TAQ_ Spread*	ρ	Number of obs.
1993	Mean	0.0273	0.0166	0.9253	34,317	0.0281	0.0175	0.9359	3,108
	Median	0.0169	0.0094			0.0167	0.0097		
1994	Mean	0.0266	0.0156	0.9113	37,339	0.0274	0.0161	0.9143	3,305
	Median	0.0174	0.0096			0.0175	0.0097		
1995	Mean	0.0320	0.0144	0.8585	37,766	0.0335	0.0151	0.8517	3,376
	Median	0.0219	0.0090			0.0224	0.0093		
1996	Mean	0.0311	0.0125	0.8267	38,772	0.0328	0.0130	0.8369	3,533
	Median	0.0208	0.0080			0.0217	0.0081		
1997	Mean	0.0316	0.0107	0.8449	40,156	0.0345	0.0111	0.8537	3,701
	Median	0.0194	0.0065			0.0200	0.0068		
1998	Mean	0.0332	0.0102	0.8571	39,362	0.0352	0.0106	0.8836	3,624
	Median	0.0220	0.0062			0.0231	0.0065		
1999	Mean	0.0415	0.0121	0.8510	39,236	0.0454	0.0129	0.9041	3,674
	Median	0.0269	0.0068			0.0289	0.0071		
2000	Mean	0.0456	0.0152	0.8980	38,857	0.0492	0.0159	0.9267	3,730
	Median	0.0286	0.0076			0.0285	0.0077		
2001	Mean	0.0317	0.0137	0.9197	36,244	0.0348	0.0140	0.9080	3,488
	Median	0.0142	0.0052			0.0158	0.0052		
2002	Mean	0.0197	0.0127	0.9357	38,555	0.0215	0.0137	0.9671	3,449
	Median	0.0096	0.0041			0.0102	0.0044		
2003	Mean	0.0114	0.0085	0.9603	38,572	0.0131	0.0096	0.9461	3,434
	Median	0.0054	0.0030			0.0056	0.0032		
2004	Mean	0.0060	0.0057	0.9531	39,952	0.0065	0.0062	0.9588	3,541
	Median	0.0027	0.0022			0.0028	0.0023		
2005	Mean	0.0053	0.0059	0.8956	41,181	0.0057	0.0065	0.9420	3,660
	Median	0.0024	0.0019			0.0026	0.0021		
2006	Mean	0.0043	0.0052	0.9065	41,235	0.0045	0.0054	0.9185	3,735
	Median	0.0018	0.0016			0.0020	0.0017		
2007	Mean	0.0045	0.0049	0.9242	43,322	0.0049	0.0056	0.9383	4,000
	Median	0.0020	0.0017			0.0021	0.0020		
2008	Mean	0.0092	0.0085	0.9406	42,416	0.0100	0.0093	0.9518	3,747
	Median	0.0030	0.0028			0.0038	0.0037		
2009	Mean	0.0103	0.0091	0.9250	35,821	0.0115	0.0102	0.9623	3,140
	Median	0.0035	0.0034			0.0040	0.0039		
	t-stat		13.9***				4.16***		
	z-stat		5.73***				1.68*		

Panel B: NASDAQ stocks

		Monthly data				Yearly data			
Year		*CRSP_ Spread*	*TAQ_ Spread*	ρ	Number of obs.	*CRSP_ Spread*	*TAQ_ Spread*	ρ	Number of obs.
1993	Mean	0.0643	0.0453	0.9193	43,891	0.0728	0.0509	0.9395	4,603
	Median	0.0489	0.0346			0.0545	0.0385		
1994	Mean	0.0635	0.0446	0.9296	50,306	0.0716	0.0501	0.9536	5,195
	Median	0.0491	0.0343			0.0540	0.0384		
1995	Mean	0.0590	0.0418	0.9356	53,109	0.0675	0.0463	0.9603	5,453
	Median	0.0434	0.0307			0.0469	0.0337		
1996	Mean	0.0544	0.0381	0.9318	59,936	0.0587	0.0411	0.9719	5,979
	Median	0.0399	0.0281			0.0429	0.0305		
1997	Mean	0.0493	0.0357	0.9472	63,812	0.0528	0.0383	0.9743	6,115
	Median	0.0350	0.0256			0.0369	0.0272		
1998	Mean	0.0429	0.0362	0.9729	62,715	0.0458	0.0388	0.9884	5,916
	Median	0.0294	0.0250			0.0318	0.0270		
1999	Mean	0.0351	0.0302	0.9597	57,606	0.0367	0.0319	0.9824	5,774
	Median	0.0240	0.0212			0.0249	0.0221		
2000	Mean	0.0337	0.0300	0.9568	57,058	0.0341	0.0305	0.9801	5,516
	Median	0.0225	0.0204			0.0241	0.0220		
2001	Mean	0.0337	0.0299	0.9551	51,101	0.0372	0.0330	0.9791	4,800
	Median	0.0215	0.0195			0.0244	0.0222		
2002	Mean	0.0299	0.0291	0.9611	45,640	0.0332	0.0310	0.9804	4,186
	Median	0.0171	0.0174			0.0191	0.0193		
2003	Mean	0.0184	0.0199	0.9518	40,889	0.0210	0.0223	0.9717	3,709
	Median	0.0099	0.0124			0.0113	0.0142		
2004	Mean	0.0115	0.0126	0.9550	38,893	0.0120	0.0130	0.9646	3,520
	Median	0.0059	0.0076			0.0064	0.0080		
2005	Mean	0.0099	0.0107	0.9500	38,057	0.0106	0.0112	0.9715	3,470
	Median	0.0044	0.0055			0.0048	0.0060		
2006	Mean	0.0078	0.0082	0.9574	37,871	0.0082	0.0085	0.9728	3,427
	Median	0.0034	0.0039			0.0038	0.0041		
2007	Mean	0.0080	0.0081	0.9621	36,647	0.0088	0.0085	0.9605	3,413
	Median	0.0035	0.0032			0.0039	0.0035		
2008	Mean	0.0230	0.0169	0.9556	36,842	0.0251	0.0180	0.9805	3,376
	Median	0.0069	0.0064			0.0096	0.0079		
2009	Mean	0.0255	0.0173	0.9657	34,469	0.0291	0.0192	0.9821	3,054
	Median	0.0069	0.0067			0.0092	0.0086		
	t-stat		12.7***				3.78***		
	z-stat		1.89*				1.01		

Panel C: Quintiles by firm size (market capitalization)

		Monthly data				Yearly data			
Size quintile		*CRSP_ Spread*	*TAQ_ Spread*	ρ	Number of obs.	*CRSP_ Spread*	*TAQ_ Spread*	ρ	Number of obs.
NYSE/ AMEX									
1 (Smallest)	Mean	0.0512	0.0302	0.8687	132,493	0.0585	0.0324	0.8872	12,022
	Median	0.0314	0.0195			0.0347	0.0213		
2	Mean	0.0225	0.0107	0.7644	132,610	0.0232	0.0111	0.8332	12,032
	Median	0.0144	0.0085			0.0155	0.0090		
3	Mean	0.0153	0.0060	0.6838	132,618	0.0158	0.0063	0.7486	12,030
	Median	0.0107	0.0049			0.0118	0.0053		
4	Mean	0.0111	0.0038	0.6521	132,611	0.0115	0.0039	0.7079	12,032
	Median	0.0083	0.0029			0.0095	0.0031		
5 (Largest)	Mean	0.0074	0.0021	0.6553	132,529	0.0076	0.0021	0.6815	12,025
	Median	0.0056	0.0015			0.0064	0.0016		
NASDAQ									
1 (Smallest)	Mean	0.0782	0.0607	0.9011	161,669	0.0886	0.0681	0.9434	15,479
	Median	0.0630	0.0501			0.0715	0.0573		
2	Mean	0.0453	0.0364	0.9080	161,781	0.0502	0.0397	0.9587	15,491
	Median	0.0380	0.0317			0.0429	0.0356		
3	Mean	0.0279	0.0226	0.9011	161,786	0.0311	0.0248	0.9455	15,484
	Median	0.0228	0.0199			0.0264	0.0224		
4	Mean	0.0180	0.0147	0.8778	161,782	0.0203	0.0163	0.9455	15,491
	Median	0.0133	0.0124			0.0160	0.0143		
5 (Largest)	Mean	0.0092	0.0076	0.8679	161,705	0.0106	0.0086	0.9161	15,482
	Median	0.0054	0.0055			0.0067	0.0066		

Table 3
Cross-sectional correlations between the TAQ spread and various low-frequency liquidity measures

We calculate the monthly cross-sectional correlation coefficients between *TAQ_Spread* and each low-frequency liquidity measure for NYSE/AMEX stocks and NASDAQ stocks, respectively, and obtain their mean values for the entire study period, 1993-2009, as well as for three sub-periods 1993-1996, 1997-2000, and 2001-2009. This table presents the mean cross-sectional correlation coefficients. We also use the *t*-test to compare the correlation coefficients between *TAQ_Spread* and each of the low-frequency liquidity measures. We put *, **, or *** on the correlation coefficient if it is significantly greater than all other correlation coefficients at the 10%, 5%, or 1% level in each period. If such correlation coefficient does not exist, we put † on the highest correlation coefficient and ‡ on the correlation coefficients that are not significantly different from the highest correlation coefficient at the 5% level according to the *t*-test. In the *t*-test, we adjust standard errors for autocorrelation with a Newey-West correction using four lags.

	CRSP_ Spread	*Roll_ Spread*	*Eff_Tick*	*Eff_Tick2*	*Gibbs*	*LOT*	*Zeros*	*Zeros2*	*Illiquidity*	*PS*
NYSE/AMEX										
Whole period	0.9020***	0.4785	0.6717	0.6376	0.7055	0.6374	0.4564	0.4231	0.7097	-0.0189
1993-1996	0.8805†	0.5190	0.5657	0.5493	0.8790‡	0.7557	0.3996	0.3400	0.7432	-0.1107
1997-2000	0.8628***	0.4988	0.6413	0.6251	0.7718	0.6408	0.4695	0.4384	0.6987	0.0392
2001-2009	0.9290***	0.4514	0.7323	0.6825	0.5990	0.5803	0.4758	0.4533	0.6997	-0.0040
NASDAQ										
Whole period	0.9515***	0.4598	0.5062	0.4810	0.6784	0.5547	0.3919	0.3034	0.6594	-0.0503
1993-1996	0.9298***	0.4764	0.4425	0.4227	0.8604	0.7451	0.3505	0.2132	0.5893	-0.0985
1997-2000	0.9591***	0.4658	0.5285	0.5148	0.7406	0.5621	0.4337	0.3367	0.6578	-0.0541
2001-2009	0.9571***	0.4502	0.5228	0.4903	0.5749	0.4720	0.3906	0.3262	0.6893	-0.0286

Table 4
Cross-sectional correlations between the TAQ spread and various low-frequency liquidity measures – Subsample analysis

We calculate the monthly cross-sectional correlation coefficients between *TAQ_Spread* and each low-frequency liquidity measure for stocks in each quintile and obtain their mean values for the entire study period, 1993-2009. This table presents the mean cross-sectional correlation coefficients. We also use *t*-test to compare the correlation coefficients between *TAQ_Spread* and each of the low-frequency liquidity measures. We put *, **, or *** on the correlation coefficient if it is significantly greater than all other correlation coefficients at the 10%, 5%, or 1% level in each period. If such correlation coefficient does not exist, we put † on the highest correlation coefficient and ‡ on the correlation coefficients that are not significantly different from the highest correlation coefficient at the 5% level according to the *t*-test. In the *t*-test, we adjust standard errors for autocorrelation with a Newey-West correction using four lags. Panel A shows the results when quintiles are formed according to market capitalizations, Panel B shows the results when quintiles are formed according to return volatility, and Panel C shows the results when quintiles are formed according to *TAQ_Spread*.

Panel A: Cross-sectional correlations – Market capitalization quintiles

Size quintile	*CRSP_ Spread*	*Roll_ Spread*	*Eff_Tick*	*Eff_Tick2*	*Gibbs*	*LOT*	*Zeros*	*Zeros2*	*Illiquidity*	*PS*
NYSE/AMEX										
1 (Smallest)	0.8687***	0.4509	0.5987	0.5606	0.8157	0.7350	0.3299	0.3234	0.6913	-0.0139
2	0.7644***	0.3394	0.6004	0.5701	0.5701	0.4921	0.1973	0.1873	0.4527	0.0080
3	0.6838***	0.2150	0.5820	0.5618	0.3598	0.3073	0.1914	0.1873	0.3985	-0.0149
4	0.6521***	0.1635	0.5665	0.5630	0.2564	0.2329	0.2618	0.2596	0.3956	-0.0461
5 (Largest)	0.6553***	0.1167	0.5695	0.5656	0.2159	0.2634	0.2472	0.2467	0.4765	-0.0555
NASDAQ										
1 (Smallest)	0.9011***	0.3096	0.3269	0.3040	0.6942	0.5816	0.1157	0.0661	0.6258	-0.0450
2	0.9080***	0.3484	0.2406	0.2384	0.5166	0.3731	0.0959	0.0588	0.6042	-0.0146
3	0.9011***	0.2895	0.2448	0.2432	0.3918	0.2715	0.1043	0.0703	0.5694	0.0009
4	0.8778***	0.2474	0.2874	0.2796	0.2970	0.2174	0.1257	0.0976	0.5349	-0.0192
5 (Largest)	0.8679***	0.1871	0.3423	0.3374	0.2088	0.1687	0.1900	0.1762	0.5253	-0.0289

Panel B: Cross-sectional correlations – Return volatility quintiles

Volatility quintile	*CRSP_ Spread*	*Roll_ Spread*	*Eff_Tick*	*Eff_Tick2*	*Gibbs*	*LOT*	*Zeros*	*Zeros2*	*Illiquidity*	*PS*
NYSE/AMEX										
1 (Smallest)	0.6795***	0.2046	0.3212	0.2984	0.3728	0.1315	0.4544	0.3332	0.5584	-0.0131
2	0.7714***	0.2340	0.4387	0.4016	0.3485	0.1574	0.6241	0.5261	0.6071	-0.0036
3	0.7989***	0.2356	0.5033	0.4680	0.3368	0.2131	0.6441	0.5612	0.6325	-0.0133
4	0.8369***	0.2558	0.5549	0.5220	0.3713	0.2552	0.6636	0.5998	0.6466	-0.0102
5 (Largest)	0.9039***	0.4660	0.6863	0.6625	0.6711	0.5785	0.6352	0.5989	0.7088	-0.0137
NASDAQ										
1 (Smallest)	0.8948***	0.2164	0.2383	0.2308	0.4936	0.2591	0.4654	0.2767	0.5575	-0.1329
2	0.9151***	0.2730	0.3124	0.2982	0.4697	0.2116	0.5677	0.4092	0.5745	-0.1271
3	0.9238***	0.3094	0.3641	0.3571	0.4733	0.2301	0.5604	0.4280	0.5937	-0.0950
4	0.9271***	0.3421	0.4010	0.3914	0.4863	0.2427	0.5391	0.4193	0.6176	-0.0651
5 (Largest)	0.9401***	0.4129	0.4727	0.4647	0.5826	0.3929	0.4559	0.3495	0.6608	-0.0337

Panel C: Cross-sectional correlations – *TAQ_Spread* quintiles

TAQ_Spread quintile	*CRSP_ Spread*	*Roll_ Spread*	*Eff_Tick*	*Eff_Tick2*	*Gibbs*	*LOT*	*Zeros*	*Zeros2*	*Illiquidity*	*PS*
NYSE/AMEX										
1 (Smallest)	0.4139‡	0.0702	0.4310†	0.4216‡	0.1467	0.2173	0.1255	0.1248	0.2549	-0.0349
2	0.2569***	0.0242	0.2379	0.2318	0.0342	0.0310	0.1316	0.1301	0.1773	-0.0186
3	0.2168***	0.0100	0.1697	0.1604	0.0087	0.0187	0.1218	0.1196	0.1249	-0.0113
4	0.3551***	00790	0.1784	0.1632	0.1415	0.1005	0.0781	0.0607	0.2573	-0.0137
5 (Largest)	0.8368***	0.4228	0.5809	0.5433	0.7512	0.6404	0.3071	0.2962	0.6794	-0.0143
NASDAQ										
1 (Smallest)	0.5799***	0.0715	0.2999	0.2904	0.1434	0.1486	0.1167	0.1272	0.1694	-0.0051
2	0.5097***	0.0683	0.1717	0.1639	0.0983	0.0733	0.0970	0.0925	0.1808	-0.0141
3	0.4934***	0.0698	0.1426	0.1304	0.1377	0.0883	0.0918	0.0842	0.2091	-0.0131
4	0.5713***	0.1165	0.1426	0.1322	0.2192	0.1285	0.0910	0.0761	0.2824	-0.0061
5 (Largest)	0.8619***	0.2508	0.2790	0.2649	0.6132	0.4747	0.1558	0.0311	0.6120	-0.0524

Table 5
Time-series correlations between the TAQ spread and various low-frequency liquidity measures

We calculate the monthly cross-sectional mean values of each liquidity measure for NYSE/AMEX stocks and NASDAQ stocks, respectively. Then we calculate the time-series correlation coefficient between the monthly mean TAQ spreads and the monthly mean values of each low-frequency liquidity measure for our entire study period, 1993-2009, as well as for the three sub-periods 1993-1996, 1997-2000, and 2001-2009. This table presents the time-series correlation coefficients. We use Fisher's *Z*-test to compare the correlation coefficients between *TAQ_Spread* and each of the low-frequency liquidity measures. We put *, **, or *** on the correlation coefficient if it is significantly greater than all other correlation coefficients at the 10%, 5%, or 1% level in each period. If such correlation coefficient does not exist, we put † on the highest correlation coefficient and ‡ on the correlation coefficients that are not significantly different from the highest correlation coefficient at the 5% level according to Fisher's *Z*-test.

	CRSP_ Spread	*Roll_ Spread*	*Eff_Tick*	*Eff_Tick2*	*Gibbs*	*LOT*	*Zeros*	*Zeros2*	*Illiquidity*	*PS*
NYSE/AMEX										
Whole period	0.7906‡	0.6139	0.8411†	0.8243‡	0.6675	0.4833	0.7307	0.7168	0.5744	-0.2985
1993-1996	-0.5073	0.8329	0.8602‡	0.8648‡	0.8843‡	0.8365	0.3681	0.3801	0.9221†	-0.4620
1997-2000	0.8961†	0.8694‡	0.6570	0.5796	0.8634‡	0.7671	0.1829	0.1326	0.7837	-0.0861
2001-2009	0.9491‡	0.7438	0.9363‡	0.9634†	0.7852	0.7429	0.5208	0.4511	0.8904	-0.2852
NASDAQ										
Whole period	0.9738***	0.9282	0.9186	0.9070	0.8560	0.6753	0.9029	0.8925	0.6250	-0.3775
1993-1996	0.9700***	0.8964	0.8264	0.8532	0.9341	0.6935	0.8288	0.8108	0.8020	-0.3476
1997-2000	0.9126***	0.7644	0.8340	0.7949	0.4273	0.2379	0.4689	0.4491	0.7550	0.1118
2001-2009	0.9342‡	0.8791	0.9460†	0.9397‡	0.8788	0.8356	0.7282	0.6845	0.9186	-0.7042

Table 6
Time-series correlations between the TAQ spread and various low-frequency liquidity measures – Subsample analysis

We calculate the monthly cross-sectional mean values of each liquidity measure for stocks in each quintile. Then we calculate the time-series correlation coefficient between the monthly mean TAQ spreads and the monthly mean values of each low-frequency liquidity measure for our entire study period, 1993-2009. This table presents the time-series correlation coefficients. We use Fisher's *Z*-test to compare the correlation coefficients between *TAQ_Spread* and each of the low-frequency liquidity measures. We put *, **, or *** on the correlation coefficient if it is significantly greater than all other correlation coefficients at the 10%, 5%, or 1% level in each period. If such correlation coefficient does not exist, we put † on the highest correlation coefficient and ‡ on the correlation coefficients that are not significantly different from the highest correlation coefficient at the 5% level according to Fisher's *Z*-test. Panel A shows the results when quintiles are formed according to market capitalizations, Panel B shows the results when quintiles are formed according to return volatility, and Panel C shows the results when quintiles are formed according to *TAQ_Spread*.

Panel A. Time-series correlations – Market capitalization quintiles

Size quintile	*CRSP_ Spread*	*Roll_ Spread*	*Eff_Tick*	*Eff_Tick2*	*Gibbs*	*LOT*	*Zeros*	*Zeros2*	*Illiquidity*	*PS*
NYSE/AMEX										
1 (Smallest)	0.6974	0.9494†	0.6764	0.6428	0.9362‡	0.8202	0.5680	0.5169	0.7600	-0.2903
2	0.8189	0.5709	0.8745†	0.8624‡	0.6644	0.4387	0.7778	0.7728	0.7444	-0.0371
3	0.8637	0.0801	0.9339†	0.9325‡	0.3343	0.1655	0.8951	0.8939	0.8644	-0.0445
4	0.8722	-0.0932	0.9639‡	0.9642†	0.1352	0.0381	0.9310	0.9306	0.8831	-0.0290
5 (Largest)	0.7864	-0.0665	0.9624‡	0.9633†	0.1289	0.0790	0.9091	0.9089	0.6806	-0.0327
NASDAQ										
1 (Smallest)	0.9580‡	0.9647‡	0.8602	0.8493	0.9718†	0.8481	0.8537	0.8256	0.6625	-0.4960
2	0.9665†	0.9581‡	0.8980	0.8857	0.9312	0.7941	0.8830	0.8732	0.7983	0.1985
3	0.9806***	0.8946	0.9293	0.9198	0.8168	0.5704	0.9306	0.9286	0.8912	0.2516
4	0.9856***	0.7094	0.9558	0.9536	0.6243	0.3657	0.9479	0.9479	0.9352	0.2512
5 (Largest)	0.9875†	0.2168	0.9775‡	0.9777‡	0.3026	0.1972	0.9466	0.9465	0.9240	0.1796

Panel B. Time-series correlations – Volatility quintiles

Volatility quintile	*CRSP_ Spread*	*Roll_ Spread*	*Eff_Tick*	*Eff_Tick2*	*Gibbs*	*LOT*	*Zeros*	*Zeros2*	*Illiquidity*	*PS*
NYSE/AMEX										
1 (Smallest)	0.7341	0.8052‡	0.7842	0.7670	0.8491†	0.6740	0.7212	0.7094	0.6249	-0.0226
2	0.8185	0.5463	0.8662†	0.8574‡	0.7062	0.4959	0.7744	0.7602	0.5359	0.0397
3	0.8132‡	0.4508	0.8582†	0.8511‡	0.6343	0.4862	0.7563	0.7477	0.5139	-0.0392
4	0.8130‡	0.4341	0.8324†	0.8235‡	0.6170	0.4795	0.7279	0.7187	0.5791	-0.1227
5 (Largest)	0.7930‡	0.7221	0.7953†	0.7782‡	0.6874	0.4839	0.6979	0.6838	0.6734	-0.3484
NASDAQ										
1 (Smallest)	0.9692***	0.8232	0.9150	0.8907	0.8241	0.5982	0.8917	0.8723	0.5486	-0.2353
2	0.9743***	0.8721	0.9261	0.9136	0.8469	0.6555	0.9008	0.8899	0.5526	-0.2208
3	0.9732***	0.8643	0.9137	0.8997	0.8301	0.6505	0.8943	0.8878	0.5308	-0.2047
4	0.9722***	0.9047	0.9125	0.9022	0.8332	0.6330	0.8990	0.8934	0.6061	-0.1990
5 (Largest)	0.9742**	0.9595	0.8996	0.8927	0.8957	0.6921	0.9001	0.8931	0.6927	-0.3883

Panel C. Time-series correlations – *TAQ_Spread* quintiles

TAQ_Spread quintile	*CRSP_ Spread*	*Roll_ Spread*	*Eff_Tick*	*Eff_Tick2*	*Gibbs*	*LOT*	*Zeros*	*Zeros2*	*Illiquidity*	*PS*
NYSE/AMEX										
1 (Smallest)	0.8050	-0.1366	0.9652†	0.9650‡	0.1130	0.0913	0.9132	0.9131	0.3164	-0.0008
2	0.8600	-0.1883	0.9602‡	0.9605†	0.0008	0.0451	0.9376	0.9371	0.9143	-0.0528
3	0.8398	0.0318	0.9445†	0.9115‡	0.2179	0.0664	0.9238‡	0.9222‡	0.8966	0.0242
4	0.8211	0.6766	0.9017†	0.8899‡	0.6864	0.4636	0.8212	0.8095	0.8722‡	0.0939
5 (Largest)	0.7291	0.9519†	0.6930	0.6630	0.9393‡	0.8285	0.5676	0.8569	0.7359	-0.2990
NASDAQ										
1 (Smallest)	0.9770†	0.1531	0.9666‡	0.9708‡	0.2447	0.1735	0.9553	0.9564	0.7180	-0.3269
2	0.9838***	0.6270	0.9585	0.9570	0.5403	0.2901	0.9596	0.9584	0.8925	-0.0142
3	0.9813***	0.8757	0.9402	0.9321	0.7449	0.4344	0.9418	0.9361	0.8815	0.1424
4	0.9726†	0.9616‡	0.9142	0.8999	0.8868	0.6786	0.9069	0.8926	0.8630	0.1405
5 (Largest)	0.9625‡	0.9634‡	0.8617	0.8487	0.9725†	0.9063	0.8592	0.8292	0.6618	-0.4723