Illiquidity as a Priced Factor: Evidence from Intradaily Data

Sahn-Wook Huh[a]

[Current Version: February 15, 2009]

[a]Faculty of Business, Brock University, St. Catharines, Ontario, Canada L2S 3A1. E-mail: shuh@brocku.ca.

I am deeply indebted to Avanidhar Subrahmanyam, Michael Brennan, and Tarun Chordia for their detailed guidance and insightful suggestions, which have made this study possible. I also thank Kee H. Chung, Kenneth Kim, Igor Kozhanov, Joseph Ogden, Maureen O'Hara, Unyong Pyo, Inho Suk, and seminar participants at the State University of New York - Buffalo and the 3rd IIROC Conference on Market Structure and Market Integrity, for valuable feedback. Research assistance was ably provided by Mi-Ae Kim. I gratefully acknowledge the generous financial support from the Social Sciences and Humanities Research Council of Canada (SSHRC) and the Faculty of Business at Brock University. All errors are solely the author's responsibility.

Abstract

Illiquidity as a Priced Factor: Evidence from Intradaily Data

A number of proxies for illiquidity have been proposed in the literature that relates trading costs to asset prices. However, some of the illiquidity measures provide equivocal relations to returns. Other measures conceal important dynamics underlying high-frequency data because they are constructed from daily or lower frequency databases. In this study, we adopt a direct and intuitive approach to estimating illiquidity. Specifically, we estimate a set of price-impact parameters based on four different models using the intradaily order flows processed via the Lee and Ready (1991) algorithm from the tick-by-tick databases for NYSE stocks over the past 23 years. Our empirical results provide strong evidence that illiquidity measured by the price-impact parameters is priced in the cross-section of stock returns, even after controlling for risk factors, firm characteristics, and other illiquidity proxies prevalent in the literature. Consistently high levels of statistical significance also suggest that the price-impact parameters estimated using the intradaily order flows are more reliable proxies for illiquidity.

Do investors require higher returns from less liquid securities? This has been an enduring question in financial economics. In a seminal paper, Amihud and Mendelson (1986) provide evidence that stock returns include a significant premium for the quoted bid-ask spread. Since that study, Brennan and Subrahmanyam (1996), Brennan, Chordia, and Subrahmanyam (1998), Jacoby, Fowler, and Gottesman (2000), Jones (2002), and Amihud (2002) all elaborate upon the role of (il)liquidity as a determinant of returns. Pástor and Stambaugh (2003) and Acharya and Pedersen (2005) relate liquidity risk to expected stock returns. Eisfeldt (2004) associates liquidity with the real sector and finds that productivity, by affecting income, feeds into liquidity. Johnson (2005) models liquidity as arising from the price discounts demanded by risk-averse agents to change their optimal portfolio holdings. Recently, Chordia, Huh, and Subrahmanyam (2008) provide evidence that theory-based estimates of illiquidity are priced in the cross-section of expected stock returns.

As we realize in the literature, an important issue in studies that relate illiquidity to asset prices is how to measure illiquidity. Other than direct empirical measurements of illiquidity by the bid-ask spread, the approach taken in the literature has often been to employ empirical arguments and econometric techniques to measure illiquidity. For example, Amihud (2002) proposes the ratio of absolute return to dollar trading volume as a measure of illiquidity. Brennan and Subrahmanyam (1996) suggest measuring illiquidity by the relation between price changes and order flows. Pástor and Stambaugh (2003) measure illiquidity by the extent to which returns reverse upon high volume, an approach based on the notion that such a reversal captures inventory-based price pressures. Hasbrouck (2005) provides a comprehensive set of measures, including the Roll (1984) measure as one of the CRSP-based proxies for a TAQ-based effective cost [because dealing with high-frequency (TAQ) data is hard]. An exception is a study by Chordia, Huh, and Subrahmanyam (2008), who turn to theory in order to derive closed-form expressions for Kyle lambdas, for which plausible empirical proxies can easily be devised from low frequency databases.

There is no denying that those measures have added considerably to our understanding of illiquidity, but there are some concerns. First, some of the measures have not yielded unambiguously consistent results. To cite a few, Brennan and Subrahmanyam

(1996) find a negative relation between bid-ask spreads and expected returns that is at odds with the liquidity premium argument. Eleswarapu and Reinganum (1993) show that the positive relationship between returns and bid-ask spreads occurs mainly in January, suggesting that the link between liquidity and expected returns is not pervasive. Spiegel and Wang (2005) also document that a variety of illiquidity proxies do not tend to be priced after controlling for the idiosyncratic risk. Second, many of the empirical measures have often been estimated or constructed from daily or low frequency databases, which may conceal some important dynamics that may exist in high frequency data. Third, in cases where (il)liquidity measures are estimated from high frequency databases, the coverage has not been broad or long enough because data availability is limited or estimation is computationally burdensome. As an alternative, therefore, theory-based measures have been derived and tested in Chordia, Huh, and Subrahmanyam (2008). The theoretical measures have advantage in the sense that broad and long input proxies for the measures are available at low costs. As is often the case, however, there may be a debate about the relevance of some empirical proxies used as inputs to the theoretical measures.

It has been recognized that a primary cause of illiquidity in financial markets is the adverse selection which is induced by privately informed traders. Many large trades often occur outside the spread and many small trades occur within the spread (Lee, 1993). Moreover, Glosten (1989), Kyle (1985), Easley and O'Hara (1987), and Glosten and Harris (1988) suggest that the illiquidity effect of asymmetric information are most likely to be captured by the price-impact of trades, or the variable component of trading costs. These facts imply that the quoted bid-ask spread may be a noisy measure of illiquidity at best. Intuitively, to examine whether illiquidity due to information asymmetry is a priced factor, it seems best to estimate the price-impact of trades and relate it to returns. However, processing intradaily order flows and estimating the price impact have been a computationally arduous task, because billions of transactions should be classified into buyer- or seller-initiated trades. That is why most studies that attempt to use order flows or price-impact parameters focus on narrow samples of stocks or short periods of time. For instance, Brennan and Subrahmanyam (1996) estimate price-impact parameters using only two-year data from the Institute for the Study of Securities Markets. But it may not be prudent to draw a general conclusion from a study that uses the two-year data, as Merton (1980) suggests that the accuracy in estimation depends on the length of data, not the sampling frequency.

Owing to improvements in computer technology, it has become possible to estimate price-impact parameters (interchangeably "lambdas")[1] for a longer period of time as well as for a broader sample of stocks, although the computing is still onerous. In this paper, we put together diverse empirical techniques from market microstructure and asset pricing research. Specifically, we adopt an intuitive and direct approach to measuring illiquidity and relate the measures to asset returns, thereby providing strong evidence in the empirical illiquidity-return relations. For this purpose, we estimate different types of price-impact parameters based on four different models using intradaily order flow data for a comprehensive set of NYSE-listed stocks over the past 23 years (276 months) from January 1983 to December 2005. To the best of our knowledge, this study uses the longest time-series of high frequency transactions data for NYSE stocks in order to process intradaily order flows and estimate a set of price-impact parameters. Order flows are processed via the Lee and Ready (1991) algorithm. To reduce the error-in-variable problem, we also adjust returns against the risk of the Fama and French (1993) factors along the lines of Brennan, Chordia, and Subrahmanyam (1998).

Our contribution in this study is that we first construct a database of marketwide intradaily order flows and the four different types of price-impact parameters in a long time series. By examining the time-series behavior of the price-impact parameters, we find a declining trend in the measures over time, which mirrors the behavior of other illiquidity proxies, such as bid-ask spreads (Jones, 2002). Then, we explore whether these lambdas as proxies for illiquidity caused by asymmetric information are priced in the cross-section of stock returns. After controlling for known characteristics such as book-to-market equity and momentum as well as for known sources of risk such as the Fama and French (1993) factors, we provide strong evidence that illiquidity measured by the four different types of lambdas is priced in the cross-section of stock returns. We check the robustness of our findings by using portfolio average lambdas and quote mid-point returns. In addition, we run a "horse race" with other commonly used illiquidity measures, demonstrating that our illiquidity measures are priced even after accounting for

[1]Throughout this paper, a *lambda* means specifically the *variable permanent* cost portion only, when the estimation models (Glosten and Harris, 1988; Foster and Viswanathan, 1993; and Sadka, 2006) decompose trading costs into several elements. For details, see Section I.

the effects of other competing illiquidity measures. These findings suggest that with all the computational difficulties, the price-impact parameters estimated using the intradaily order flows are among the most reliable proxies for illiquidity in stock markets.

The remainder of this paper is organized as follows. In Section I, we present the algorithm to process order flows as well as the four models to estimate price-impact parameters using the intradaily order flow data. Section II describes the methodology for analyses. Section III outlines data sources, definitions, descriptive statistics, and data adjustments. Section IV discusses the empirical results and robustness checks. In Section V, we compare the effects of the price-impact parameters with those of other alternative illiquidity measures. Section VI concludes.

I. Estimation of Price-Impact Parameters

A. Lee and Ready's (1991) Algorithm

To estimate price-impact parameters, which will be used as our illiquidity measures, from monthly regressions, we first should obtain intradaily order flows. Before describing the models on which the estimation of the price-impact parameters are based, we briefly explain how the intradaily order flows (signed volume) are processed from trade and quote databases.

Given that not all trades occur at the bid or ask price, we classify each trade as a buyer- or seller-initiated trade according to the Lee and Ready (1991) algorithm using trades and quotes data from the Institute for the Study of Securities Markets (ISSM: 1983-1992) and the NYSE Trades and Automated Quotations (TAQ: 1993-2005) databases. In order to match trades and quotes, any quote less than five seconds prior to the trade is ignored and the first one at least *five* seconds prior to the trade is retained for the years 1983 to 1998. Based on feedback from microstructure scholars, who indicate that timing differences in recording trades and quotes have dramatically declined in recent years, this typical five-second delay rule in matching is not imposed for the last seven years. Instead, the quote immediately prior to each transaction (i.e., the quote closest in time to the transaction, with a time stamp of *two* seconds or more before the transaction) is retained for the last seven years (1999-2005).

Then the transactions data are signed as follows. If a trade occurs above the prevailing quote mid-point, it is regarded as buyer-initiated and *vice versa*. If a trade occurs exactly at the quote mid-point, we discard the trade. Admittedly, there may be some signing errors in the process. In this sense, the resulting order flows are *estimates*. However, Lee and Radhakrishna (2000) and Odders-White (2000) show that the algorithm is accurate enough not to pose any serious problem.

B. Models for Estimating Price-Impact Parameters

Given the intradaily order flow data processed as above, we now present some methods and models used to estimate price-impact parameters in four different ways as follows.

B.1. Estimation Based on Kyle (1985)

The first one of the four price-impact parameters (lambdas) is estimated based on a simple model of Kyle (1985). The Kyle (1985) model suggests that

$$\Delta P_{i,t,m} = \lambda^K_{i,m} S_{i,t,m} V_{i,t,m} + \epsilon_{i,t,m}, \tag{1}$$

where $\Delta P_{i,t,m}$ is a price change (in stock i at time t in month m), $S_{i,t,m}$ is the sign of a trade ($S = +1$ if the trade is buyer-initiated, and $S = -1$ if it is seller-initiated), and $V_{i,t,m}$ is volume (share volume or dollar volume) of the trade. $S_{i,t,m}V_{i,t,m}$ is now *signed volume*, which is often called *order flows*. Our price-impact parameter, $\lambda^K_{i,m}$ (Kyle lambda), is estimated each month for each stock by running the time-series regressions in Eq.(1) (with a constant term) using the intradaily order flows available within month m (m =1983:01 to 2005:12).

B.2. Estimation Based on Glosten and Harris (1988)

Glosten and Harris (1988) decompose trading costs into four components: i) fixed permanent cost (denoted as $\overline{\lambda}$), ii) variable permanent cost (λ), iii) fixed transitory cost ($\overline{\varphi}$), and iv) variable transitory cost (φ). The first two components are due to adverse selection or asymmetric information, while the last two are due to inventory holding costs, clearing fees, and/or monopoly power.

Let μ_t denote the expected value of a security (conditional on the information set at time t) for a market maker who observes only the order flows (S_tV_t) and the public information signal (ξ_t). Then, models of price formation such as Kyle (1985) and Admati and Pfleiderer (1988) imply that μ_t evolves as follows:

$$\mu_t = \mu_{t-1} + \lambda S_t V_t + \xi_t. \tag{2}$$

Glosten and Harris (1988) show evidence that the fixed permanent cost and the variable transitory cost are negligible in their sample (14 months from December 1981 to January 1983): i.e., $\overline{\lambda} = \varphi = 0$. In the estimation, we reflect their finding and assume competitive risk-neutral market makers. Given the sign (S_t) of each trade, we can write the observed security price, P_t, as

$$P_t = \mu_t + \overline{\varphi} S_t. \tag{3}$$

Plugging Eq.(2) into Eq.(3), we have

$$P_t = \mu_{t-1} + \lambda S_t V_t + \overline{\varphi} S_t + \xi_t. \tag{4}$$

From Eq.(3), we also know

$$P_{t-1} = \mu_{t-1} + \overline{\varphi} S_{t-1}. \tag{5}$$

If we subtract Eq.(5) from Eq.(4) and use the notations in a more specific way for our purpose, the price change, ΔP_t, is given by

$$\Delta P_{i,t,m} = \lambda_{i,m}^{GH} S_{i,t,m} V_{i,t,m} + \overline{\varphi}_{i,m}^{GH} (S_{i,t,m} - S_{i,t-1,m}) + \xi_{i,t,m}, \tag{6}$$

where $\lambda_{i,m}^{GH}$ (Glosten-Harris lambda) is our second measure of illiquidity (for stock i in month m), $\overline{\varphi}_{i,m}^{GH}$ is the fixed transitory cost, and $\xi_{i,t,m}$ is the unobservable error term. To estimate the Glosten-Harris lambda ($\lambda_{i,m}^{GH}$) each month for each stock, we run the time-series regressions as in Eq.(6) (with a constant term) using all the intradaily order flows available within month m.

B.3. Estimation Based on Foster and Viswanathan (1993)

To estimate a measure of the adverse selection component of price changes, Foster and Viswanathan (1993) use the *unexpected* order flows, instead of the raw order flows used in Eq.(6) above. This approach has an advantage because, if the order flows are autocorrelated, then part of the order flows is predictable and should not be included in measuring the information content of a trade.

Following Brennan and Subrahmanyam (1996) and Sadka (2006), we thus filter the order flows by an AR(5) process as in the following equation,

$$S_t V_t = \delta + \sum_{q=1}^{5} \kappa_q S_{t-q} V_{t-q} + \tau_t, \tag{7}$$

where τ_t is the residual from the time-series regression. We use τ_t as the unexpected order flows to estimate price-impact parameters, replacing $S_t V_t$ with τ_t in Eq.(6) as follows:

$$\Delta P_{i,t,m} = \lambda_{i,m}^{FV} \tau_{i,t,m} + \overline{\varphi}_{i,m}^{FV} (S_{i,t,m} - S_{i,t-1,m}) + \xi'_{i,t,m}, \tag{8}$$

where $\lambda_{i,m}^{FV}$ (Foster-Viswanathan lambda), as our third measure of illiquidity for stock i in month m, is now the response to the unexpected portion of the order flows, and $\overline{\varphi}_{i,m}^{FV}$ is the corresponding fixed transitory cost. To estimate $\lambda_{i,m}^{FV}$ each month for each stock, we run the time-series regressions as in Eq.(8) (with a constant term) using all the intradaily order flows available within month m.

B.4. Estimation Based on Sadka (2006)

Unlike Glosten and Harris (1988), Sadka (2006) documents that the fixed permanent cost ($\overline{\lambda}$) and the variable transitory cost (φ) do not tend to be zero. In this sense, Sadka's (2006) specification is equivalent to the full version of Glosten and Harris (1988).

Following Sadka (2006), we first estimates the unexpected order flows (τ_t), their variance (σ_τ^2), and the fitted value of order flows ($\widehat{S_t V_t}$) from Eq.(7). Then we compute the unexpected sign (π_t) of a trade using the following equation,

$$\pi_t = S_t - E_{t-1}(S_t) = S_t - \left\{1 - 2\Phi\left(-\frac{\widehat{S_t V_t}}{\sigma_\tau}\right)\right\},$$

where $\Phi(.)$ denotes the normal cumulative distribution function. Now the four components of trading costs can be estimated by the following regression:

$$\begin{aligned} \Delta P_{i,t,m} &= \overline{\lambda}^S_{i,m}(S_{i,t,m}V_{i,t,m} - S_{i,t-1,m}V_{i,t-1,m}) + \lambda^S_{i,m}\tau_{i,t,m} + \varphi^S_{i,m}\pi_{i,t,m} \\ &\quad + \overline{\varphi}^S_{i,m}(S_{i,t,m} - S_{i,t-1,m}) + \varsigma_{i,t,m}, \end{aligned} \tag{9}$$

where $\overline{\lambda}^S_{i,m}$ is the fixed permanent cost, $\lambda^S_{i,m}$ (Sadka lambda), as our fourth measure of illiquidity, is the variable permanent cost, $\varphi^S_{i,m}$ is the variable transitory cost, and $\overline{\varphi}^S_{i,m}$ is the fixed transitory cost.[2] To estimate Sadka lambda ($\lambda^S_{i,m}$) each month for each stock, we run the time-series regressions as in Eq.(9) (with a constant term) using all the intradaily order flows and other related variables available within month m. Note that when $\overline{\lambda}^S_{i,m} = \varphi^S_{i,m} = 0$, Eq.(9) is reduced to the parsimonious Foster and Viswanathan (1993) model specified in Eq.(8).

As we see above, the last three models decompose trading costs into two to four components. In those cases, given that we are more interested in the trading cost related to information asymmetry, we use the price-impact parameters that represent the variable permanent cost only (i.e., $\lambda^{GH}_{i,m}$, $\lambda^{FV}_{i,m}$, and $\lambda^S_{i,m}$), ignoring the other components of trading costs. Another issue that arises when we estimate the price-impact parameters in Eq.(1)-Eq.(9) is whether to use *dollar* order flows or *share* order flows. When running the time-series regressions, we design the programs so that the parameters are estimated in both ways, resulting in the eight different types of measures. For our main analyses, we employ the price-impact parameters estimated with dollar order flows as our primary basis, because we believe that dollar order flows can better estimate the price impact. For robustness checks, however, we also discuss later the results from using the parameters estimated with share order flows.

[2]For full derivation of Eq.(9), see Sadka (2006).

II. Methodology

For asset pricing tests, we follow the Brennan, Chordia, and Subrahmanyam (BCS) (1998) approach, which uses data on individual securities [Hasbrouck (2005) also uses this approach]. The BCS methodology is important because using portfolios could be problematic, as Roll (1977) and Lo and MacKinlay (1990) suggest. This approach not only avoids the data-snooping biases that are inherent in the portfolio-based approaches but also gets around the error-in-variable biases caused by errors in estimating factor loadings.

Now assume that returns are generated by an L-factor approximate factor model:

$$\tilde{R}_{jt} = E(\tilde{R}_{jt}) + \sum_{k=1}^{L} \beta_{jk}\tilde{f}_{kt} + \tilde{e}_{jt}, \tag{10}$$

where $\tilde{R}_{jt}$ is the return on security j at time t and $\tilde{f}_{kt}$ is the unanticipated return on the k-th factor $(k = 1, 2, ..., L)$ at time t. The exact, or equilibrium, version of the arbitrage pricing theory (APT) in which the market portfolio is well diversified with respect to the factors (Connor, 1984; Shanken, 1985, 1987) implies that the expected excess returns may be written as

$$E(\tilde{R}_{jt}) - R_{Ft} = \sum_{k=1}^{L} \theta_{kt}\beta_{jk}, \tag{11}$$

where R_{Ft} is the return on the risk-free asset and θ_{kt} is the risk premium on the k-th factor portfolio. Plugging Eq.(11) into Eq.(10), the APT implies that realized returns are given by

$$\tilde{R}_{jt} - R_{Ft} = \sum_{k=1}^{L} \beta_{jk}\tilde{F}_{kt} + \tilde{e}_{jt}, \tag{12}$$

where $\tilde{F}_{kt} \equiv \theta_{kt} + \tilde{f}_{kt}$ is the sum of the risk premium on the k-th factor portfolio and the innovation of the k-th factor.

Our goal is to test whether illiquidity (due to information asymmetry) measured by the price-impact parameters has any incremental explanatory power for returns relative to the Fama and French (FF, 1993) 3-factor benchmark after controlling for other security characteristics. For this purpose, a standard application of the Fama-MacBeth (1973) procedure would involve estimation of the following equation:

$$\tilde{R}_{jt} - R_{Ft} = c_0 + \phi\lambda^i_{jt} + \sum_{k=1}^{L} \theta_k \beta_{jkt} + \sum_{n=1}^{N} c_n Z_{njt} + \tilde{e}_{jt}, \tag{13}$$

where λ^i_{jt} ($i = K$, GH, FV, or S) is one of our illiquidity measures estimated in Section I, and a vector of control variables, Z_{njt}, is firm characteristic n ($n = 1$, ..., N) for security j in month t. Under the null hypothesis that excess returns depend only on risk as measured by β_{jk}, coefficients ϕ and c_n ($n = 1$, ..., N) will be zero. This hypothesis can be tested in principle by first estimating the factor loadings each month using the past data; conducting a cross-sectional regression for each month in which the independent variables are an illiquidity measure, factor loadings, and other non-risk characteristics; and then averaging the monthly coefficients over time and computing their standard errors. This basic Fama-MacBeth approach, however, will present a problem if the factor loadings are measured with errors.

In order to address the errors-in-variables problem, we use risk adjusted-returns as the dependent variables. Risk adjustment is made using the Fama-French (1993) three factors (interchangeably "FF3": MKT_t, SMB_t, and HML_t) in two different ways. In the first method, we compute risk-adjusted returns, $\tilde{R}^*_{jt}$, for each month as the sum of the intercept and the residual, i.e.,

$$\begin{aligned} \tilde{R}^*_{jt} &= (\tilde{R}_{jt} - R_{Ft}) - (\widehat{\beta}^*_{j1} MKT_t + \widehat{\beta}^*_{j2} SMB_t + \widehat{\beta}^*_{j3} HML_t) \\ &= \widehat{\alpha}^* + \widehat{\tilde{e}}^*_{jt}, \end{aligned} \tag{14}$$

after conducting regressions in Eq.(12) (but with a constant term α) using the *entire* sample range (from January 1983 to December 2005 for NYSE stocks) of the data.[3] We denote this risk-adjusted return ($\tilde{R}^*_{jt}$) as *FF3EXSRET1*. We also use another version of risk adjustment for robustness. In the second method, we obtain *rolling* estimates of the factor loadings, β_{jk}, for each month over the sample period for all securities using the time series of the past 60 months (at least 24 months) with Eq.(12). Given the current month's data ($\tilde{R}_{jt}$-R_{Ft}, MKT_t, SMB_t, and HML_t) and the factor loadings ($\widehat{\beta}^{**}_{jk}$)

[3]In the first method, therefore, for each stock we have only *one* set of the factor loadings ($\widehat{\beta}^*_{jk}$), estimated using the whole time-series of the data.

estimated each month for all stocks, we can compute the risk-adjusted return on each of the securities, $\tilde{R}^*_{jt}$, for each month t as follows:

$$\tilde{R}^*_{jt} = (\tilde{R}_{jt} - R_{Ft}) - (\widehat{\beta}^{**}_{j1} MKT_t + \widehat{\beta}^{**}_{j2} SMB_t + \widehat{\beta}^{**}_{j3} HML_t). \tag{15}$$

This risk-adjusted return ($\tilde{R}^*_{jt}$) is notated as *FF3EXSRET2*.

The risk-adjusted returns from Eq.(14) and Eq.(15) constitute the raw material for the estimates that we present in the following Fama-Macbeth (1973) cross-sectional regressions:

$$\tilde{R}^*_{jt} = c_{0t} + \phi_t \lambda^i_{jt} + \sum_{n=1}^{N} c_{nt} Z_{njt} + \tilde{e}'_{jt}, i = K,\ GH,\ FV,\ \text{or}\ S. \tag{16}$$

Note that the error term in Eq.(16) is different from that in Eq.(13) because the error in Eq.(16) also contains terms arising from the measurement error associated with the factor loadings.

To check whether illiquidity is priced, we report three types of statistics based on regressions in Eq.(16): the statistics based on regressions with the dependent variable in Eq.(16) being (i) risk-*unadjusted* excess returns (we call this unadjusted return EXSRET); (ii) risk-adjusted excess returns using the first method, FF3EXSRET1; and (iii) risk-adjusted excess returns using the second method, FF3EXSRET2. For our purposes, we estimate the vector of coefficients $\mathbf{c}_t = [c_{0t}\ \phi_t\ c_{1t}\ c_{2t}...c_{Nt}]'$ from Eq.(16) each month with a simple OLS regression as

$$\widehat{\mathbf{c}}_t = (\mathbf{Z}'_t\mathbf{Z}_t)^{-1}\mathbf{Z}'_t\tilde{\mathbf{R}}^{*h}_t,$$

where $h = 1$ or 2, $\mathbf{Z}_t = [\lambda^i\ Z_1\ Z_2...Z_N]'$ and $\tilde{\mathbf{R}}^{*h}_t$ is the vector of risk-adjusted excess returns based on Eq.(14) or Eq.(15). The standard Fama-MacBeth (1973) estimator is the time-series average of the monthly coefficients, and the standard error of this estimator is taken from the time series of monthly coefficient estimates, $\widehat{\mathbf{c}}_t$. Note that although factor loadings are estimated with error in Eq.(12), this error affects only the dependent variable, $\tilde{R}^{*h}_t$, as we see in Eq.(14), Eq.(15), and Eq.(16). While the factor loadings will be correlated with vector $\mathbf{Z}_t = [\lambda^i\ Z_1\ Z_2...Z_N]'$, there is no *a priori* reason to believe that the errors in the estimated loadings will be correlated with the vector $\mathbf{Z}_t$. This implies that coefficient vector $\widehat{\mathbf{c}}_t$ estimated in Eq.(16) is unbiased.

If the errors in the estimated factor loadings are correlated with the explanatory variables $\mathbf{Z}_t$, the monthly estimates of the coefficients, $\widehat{\mathbf{c}}_t$, will be correlated with the factor realizations, and thus the mean of these estimates (which is the Fama-MacBeth estimator) will be biased by an amount that depends on the factor realizations. Therefore, as a check on the robustness of our results, we also obtained a "purged" estimator for each of the explanatory variables in the regressions of FF3EXSRET1 and FF3EXSRET2: i.e., the constant term (and its t-value) from the regression of the monthly coefficients ($\widehat{\mathbf{c}}_t$) estimated in Eq.(16) on the time series of FF 3 factor realizations. This estimator, which was developed by Black, Jensen, and Scholes (1972), purges the monthly estimates of the factor-dependent component so that it is unbiased even when the errors in the factor loading estimates are correlated with vector $\mathbf{Z}_t$.

III. Data, Definitions, Descriptive Statistics, and Adjustments

For examining the impact of illiquidity on the cross-section of stock returns, we mainly use NYSE-listed firms at an intradaily, daily, and monthly frequency over the 276 months (23 years) from January 1983 to December 2005.

A. Order Flows and Price-Impact Parameters

To process order flows via the Lee and Ready (1991) algorithm and estimate the four types of price-impact parameters based on the models described in the previous section, we focus only on NYSE stocks (the case where the exchange code is "N" in trade and quote databases), because NASDAQ has different trading protocols (Atkins and Dyl, 1997). The intradaily transaction data sources are the Institute for the Study of Securities Markets (ISSM) and the NYSE Trades and Automated Quotations (TAQ). The ISSM data are available for 1983-1992 while the TAQ data are used for 1993-2005. If trades are out of sequence or they are recorded before the open or after the closing time, such trades are expunged. Quotes established before the opening of the market or after the close are also excluded. To minimize possible signing errors in processing order flows, if a trade occurs exactly at the quote mid-point, we discard the trade, as in Sadka (2006), before running regressions to estimate the price-impact parameters. We find that about 5% of the trades from the intradaily databases are transacted at the quote mid-points.[4] To survive in the sample, stocks should have at least 110 trades per month (on average 5 trades per day) for each firm.

Table I summarizes the intradaily order flow data processed in order to estimate the price-impact parameters. The total number of trades (and matched quotes) used in this study is slightly larger than two billion over the sample period (276 months), excluding trades executed exactly at the quote mid-points. By construction, the minimum number of monthly trades (and matched quotes) is 110 for each firm, given our inclusion requirements specified above. Shares of a typical firm traded 3,636.5 times on average within a month (excluding the trades executed at the quote mid-points), but we find that some firms such as Chevron Corporation (ticker symbol: CVX) and Exxon Mobil Corporation (ticker symbol: XOM) had extremely frequent trades in recent months. For instance, the number of monthly trades in shares of Chevron Corporation is as large as 173,471 in October 2005. The table also exhibits that the pooled number of firms used over the sample period (i.e., firm-month observations) is 443,453.

In Table II, we report the sign, statistical significance, and descriptive statistics for the price-impact parameters estimated using the dollar volume-based order flows as delineated in Section I. Each of the lambdas shown in Table II is defined as follows:

λ_0^K: The price-impact parameter estimated based on Kyle (1985) (multiplied by 10^6) using intradaily dollar order flows available within each month for each stock.

λ^K: The Winsorized parameter of λ_0^K at the 0.5th and 99.5th percentiles (i.e., λ_0^K's less than the 0.5 percentile value or greater than the 99.5 percentile value in each month are set equal to the 0.5th and 99.5th percentile values, respectively).

λ_0^{GH}: The price-impact parameter estimated based on Glosten and Harris (1988) (multiplied by 10^6) using intradaily dollar order flows available within each month for

[4]For example, 5.3% of the total trades in 2002 occur exactly at the quote mid-points. When a trade occurs exactly at the quote mid-point, we also try to classify the trade using an alternative method called a tick test. That is, if a trade occurs at the quote mid-point, it is signed using the previous transaction price: buyer-initiated if the sign of the last non-zero price change is positive, and *vice versa*. This method does not change our main results.

each stock.

λ^{GH}: The Winsorized parameter of λ_0^{GH} at the 0.5th and 99.5th percentiles.

λ_0^{FV}: The price-impact parameter estimated based on Foster and Viswanathan (1993) (multiplied by 10^6) using intradaily dollar order flows available within each month for each stock.

λ^{FV}: The Winsorized parameter of λ_0^{FV} at the 0.5th and 99.5th percentiles.

λ_0^{S}: The price-impact parameter estimated based on Sadka (2006) (multiplied by 10^6) using intradaily dollar order flows available within each month for each stock.

λ^{S}: The Winsorized parameter of λ_0^{S} at the 0.5th and 99.5th percentiles.

We expect that the estimated price-impact parameters will be positive and statistically significant in most cases. Panel A in Table II shows that indeed 99.8% of Kyle (1985) lambdas are positive and 97.2% of them are both positive and statistically significant at the 5% level. Given that the other three types of lambdas are estimated by the models that decompose trading costs into several components or use unexpected order flows, it is natural to observe that the proportion of positive and significant lambdas decreases. Especially, as for the lambdas estimated based on the Sadka (2006) model, which decomposes trading costs into four components and uses unexpected order flows, 93.4% of them are positive while 67.2% are both positive and significant.

Panel B in Table II reports the time-series average values of monthly means, medians, standard deviations (STD), and other descriptive statistics for our price-impact parameters. The mean values of the statistics are first computed cross-sectionally each month and then averaged in the time-series over the sample period. A noteworthy aspect is that the four non-Winsorized lambdas (λ_0^{K}, λ_0^{GH}, λ_0^{FV}, and λ_0^{S}) are highly leptokurtic as well as significantly skewed to the left. The large kurtoses of the lambdas also imply that sample distributions of the four measures exhibit many extreme observations. To alleviate the influence of extreme observations on the empirical results, Hasbrouck (1999, 2005, 2006) and Chordia, Huh, and Subrahmanyam (2008) apply square-root transformation to illiquidity measures while others prefer logarithmic transformation. However, neither transformation is feasible for our price-impact parameters, because some of them are negative as we see in Panel A. For this reason, as our illiquidity measures we use the price-impact parameters Winsorized at the 0.5th and 99.5th percentiles for the empirical analyses in the next sections. As we see in Panel B, the skewnesses and kurtoses of the corresponding Winsorized parameters (λ^K, λ^{GH}, λ^{FV}, and λ^S) are substantially reduced by the procedure. Another feature is that the magnitude of the Winsorized lambdas decreases monotonically as the estimation model decomposes trading costs into several components and/or uses unexpected order flows. Thus, the average value of λ^S is the smallest while that of λ^K is the largest out of the four illiquidity measures.

To examine the time-series behavior of our four illiquidity measures, in Figure 1 we plot the value-weighted series of the Winsorized price-impact parameters over the sample period. As we see in Figure 1(a), value-weighted λ^K of NYSE stocks remains high in the initial part of the sample period. However, it exhibits a decreasing time trend in general, suggesting that market liquidity has improved since the early 1980s. In Figures 1(b), the trend of the value-weighted λ^S is comparable to that of λ^K, but the absolute level in this measure is much lower. We observe some big spikes around the months of, for instance, the 1987 stock market crash. For brevity, we do not report the graphs for λ^{GH} and λ^{FV}, but their trends are qualitatively similar, with the levels being in between those of the two graphs plotted in Figure 1.

B. Other Definitions and Descriptive Statistics

The three dependent variables (EXSRET, FF3EXSRET1, and FF3EXSRET2) defined in Section II for the Fama and MacBeth (1973) regressions are obtained or estimated using the CRSP monthly file, and the three FF factors are available from Kenneth French's website. We also use firm characteristics in the regressions as control variables. The probable control variables and other related variables are defined as follows:

MV: The market value defined as the month-end stock price times the number of shares outstanding (in $million) in the previous month (as of month $t-1$).

SIZE: The natural logarithm of MV.

BM_w: The Winsorized value (at the 0.5th and 99.5th percentiles) of a book-to-market ratio (BM = BV/MV), where the book value (BV) is defined as common equity plus deferred taxes in $million.

BTM: The natural logarithm of BM_w. Following Fama and French (1992), we fill monthly BM_w (and hence BTM) values for July of year y to June of year $y + 1$ with the value computed using the accounting data at the end of year $y - 1$, assuming a lag of six months before the annual accounting numbers are known to investors.

MOM1: The compounded holding period return of a stock over the most recent 3 months (from month $t - 1$ to month $t - 3$).

MOM2: The compounded holding period return over the next recent 3 months (from month $t - 4$ to month $t - 6$).

MOM3: The compounded holding period return over the 3 months from month $t - 7$ to month $t - 9$.

MOM4: The compounded holding period return over the 3 months from month $t - 9$ to month $t - 12$. For each of the above four momentum variables to exist, a stock should have all three consecutive monthly returns over the corresponding three-month period.

Later in Section V, we run a horse race to compare the effects of our four illiquidity measures with those of three other alternative illiquidity measures commonly used in the literature. The alternative measures to be analyzed in our study are notated and defined as follows:

Amihud: The Winsorized (at the 0.5th and 99.5th percentiles) illiquidity measure of Amihud (2002). We estimate this measure each month as the average of $|r|/DVOL$, where r is the daily stock return and $DVOL$ is the daily dollar volume in $100,000.

Roll_Gibbs: The Winsorized (at the 0.5th and 99.5th percentiles) market risk-adjusted effective bid-ask spread of Roll (1984), estimated at an annual frequency using the Gibbs sampler. This measure was obtained from Joel Hasbrouck's website.

PS: The Winsorized (at the 0.5th and 99.5th percentiles) illiquidity measure of Pástor and Stambaugh (2003). We estimate this measure by running monthly regressions using the CRSP daily data whose transaction records are kept for at least 15 days within a month (see Section V for details).

The variables related to the book-to-market ratio are constructed using the CRSP and CRSP/Compustat Merged (CCM) files. Other firm characteristic and related variables (MV, SIZE, and MOM1-MOM4) are also extracted from the CRSP monthly file. The two alternative illiquidity measures, *Amihud* and *PS*, are estimated using the CRSP daily file. The average number of component stocks used each month in the Fama-MacBeth (1973) cross-sectional regressions for NYSE stocks is 1,578.1. In those cases where accounting variables and other data are available only on a yearly basis, we keep the relevant values constant for 12 months in the regressions.[5]

Table III reports the time-series average values of monthly means, medians, standard deviations (STD), and other descriptive statistics for characteristic variables. To obtain the values in the table, the mean value of each statistic is first computed cross-sectionally each month and then the time-series is averaged over the sample period. As we see in the table, the average market value (MV) is $3.19 billion, and the book-to-market ratio (BM_w) is 0.74 on average for NYSE stocks over the sample period. Both variables tend to be left-skewed. As for the momentum variables (MOM1-MOM4), the returns of the distant past are likely to be higher than those of the recent past.

Next, we examine the average correlation coefficients between our explanatory variables in Table IV.[6] The table shows that our four illiquidity measures are highly correlated with each other: A maximum of 99.6% between λ^{GH} and λ^{FV} and a minimum of 77.1% between λ^{K} and λ^{S}. The illiquidity measures are negatively correlated with the momentum variables, suggesting that good past price performance of a stock tends to contribute to the improvement in liquidity of that stock. The correlation coefficients between the book-to-market ratio and the illiquidity measures are positive, although not so strong.

We would expect larger firms (with greater breadth of ownership) to be more liquid than smaller ones. As we see in Table IV, indeed SIZE and the four illiquidity measures are highly negatively correlated: the grand average of the cross-sectional correlation coefficients ranges from -45% to -57%. A preliminary test shows, however, that at the individual (month or firm) level the coefficients of time-series correlation (for the whole sample period for each firm) between SIZE and the illiquidity measures are often higher than -65%, while those of cross-sectional correlation (each month) are often higher than

[5] The data series available only on a yearly basis are the variables related to the book-to-market ratio (BM_w and BTM) and the market risk-adjusted effective bid-ask spread of Roll (1984) estimated using the Gibbs sampler (*Roll_Gibbs*).

[6] To save space, we do not report the correlation coefficients between the three types of excess returns to be used as a left-hand side variable in the regressions. They are strongly correlated (coefficients greater than 94%).

−60%. We have realized that this causes a multicollinearity problem when we include SIZE in the regressions (see Section IV and Table A1 later). Therefore, SIZE (as a firm characteristic variable) will not be used in subsequent regression analyses.

C. Gallant, Rossi, and Tauchen's (1992) Adjustments

Some of our time-series are non-stationary. This creates the potential problem that the time-series average of the cross-sectional coefficients as in Fama and MacBeth (1973) may not converge to the population estimates. According to results from the Dickey-Fuller unit-root tests and our own intuition, the obvious candidates for non-stationarity are our illiquidity measures (λ^K, λ^{GH}, λ^{FV}, and λ^S), the three alternative illiquidity measures ($Amihud$, $Roll\_Gibbs$, and PS), and some of the firm characteristic variables (BTM and $SIZE$). To eliminate non-stationarity, we adjust these data series in two steps along the lines of Gallant, Rossi, and Tauchen (1992) before conducting cross-sectional regressions. Calendar effects and trends are removed from the means and the variances of the above data series over the sample period for each of all the component stocks. As adjustment regressors, we use eleven dummy variables for months (January-November) of the year as well as the linear and quadratic time-trend variables (t and t^2).

In the first stage, we regress each of the series to be adjusted on the set of the adjustment regressors for each firm over the sample period as in the mean equation:

$$\kappa = x'\psi + \xi, \tag{17}$$

where κ represents one of the above series to be adjusted, and x is a vector of ones and the adjustment regressors (11 monthly dummies, t, and t^2). In the second stage, we take the residuals from the mean equation to construct the following variance equation:

$$\log(\xi^2) = x'\theta + \epsilon. \tag{18}$$

This regression standardizes the residuals from the above mean equation. Then we finally can obtain the adjusted series for each firm by the following linear transformation:

$$\kappa_{adj} = a + b\{\widehat{\xi}/\exp(x'\theta/2)\}, \tag{19}$$

where a and b are chosen so that the sample means and variances of κ and κ_{adj} are the same. This linear transformation makes sure that the units of adjusted and unadjusted series are equivalent, facilitating interpretation of our empirical results in the next sections. After the Gallant, Rossi, and Tauchen (1992) (GRT)-adjustments, the Dickey-Fuller tests show no evidence of a unit root in the vast majority of the component stocks over the sample period.[7]

IV. Empirical Results

A. *Features of the Portfolios Formed on Illiquidity and Firm Size*

Before moving on to regression analyses, we report the average values of firm size, illiquidity, and monthly returns for the 25 portfolios formed by sorting on illiquidity (measured by λ^{GH}) and firm size (measured by MV). For this purpose, each month we first sort sample stocks by λ^{GH} in ascending order and split them into five portfolios with the equal number of stocks. Then, each of the five portfolios is again sorted by firm size (MV) and split into five portfolios, resulting in the 25 portfolios.[8] Next, the cross-sectional average values of firm size, illiquidity, and market value (MV)-weighted returns are computed each month for each of the 25 portfolios, and then the time-series averages of the three variables are reported in Table V.

It is reasonable to observe in Panel A that average firm size (within a given firm-size group) is decreasing as illiquidity increases. Panel B also shows that average illiquidity (within a given λ^{GH} group) is mostly decreasing as firm size increases (except for the most liquid group).

We see in the upper part of Panel C that for a given firm-size group the value-weighted average return tends to increase (although not monotonic) with illiquidity. The

[7]For each GRT-adjusted variable, the unit root hypothesis is rejected for more than 95% of the sample stocks. Specific percentages are available on request. When the GRT-adjusted variables are used in the cross-sectional regressions over the next two sections, such variables will be indicated by superscript "a".

[8]In forming the 25 portfolios, we use λ^{GH} as a representative illiquidity measure because the Glosten and Harris (1988) model decomposes trading costs into several components. However, using any of the four illiquidity measures leads to similar results.

t-values demonstrate that monthly portfolio returns are all significantly different from zero. In particular, the bottom part of Panel C shows that the average returns in all of the five differential ($Illiqid - Liquid$) portfolios are positive and four cases of them are statistically different from zero at the 5% level. This suggests that illiquidity is priced for NYSE stocks and this illiquidity pricing is not driven completely by a few small firms. Another aspect is that in illiquidity quintile 5, the return shows a decreasing pattern (more precisely, a flipped J-curve) as firm size increases. For illiquidity quintiles 1-4, however, the return in a given illiquidity group is likely to increase with firm size, which is consistent with the result documented by Brennan and Subrahmanyam (1996).[9]

To get a feel about the impact of illiquidity on returns, we also report in Table VI the intercept (notated as $FF3$ *Alpha*) and t-statistic from the time-series regression of the portfolio return (in excess of the one-month T-bill rate) on the Fama-French 3 factors using the time series of the value-weighted portfolio return in each of the 25 portfolios. The upper part of Panel A shows that while the value-weighted portfolio returns over and above those predicted by the FF 3-factor model (i.e., $FF3$ *Alpha*) for the two smaller-sized groups tend to be negative, the abnormal returns for the other three size groups (firm-size quintiles 3-5) or the most illiquid group (illiquidity quintile 5) are positive and mostly statistically different from zero at 5%. For a given firm-size group, the abnormal return ($FF3$ *Alpha*) tends to increase (although not monotonic) with illiquidity.

In the lower part of Panel A, we report the five $FF3$ *Alphas* from the time-series regressions of the return differences [between the most illiquid portfolio (*Illiquid*) and the most liquid portfolio (*Liquid*) within a given firm-size group] on the FF 3 factors, together with their t-statistics. The t-statistics indicate that the abnormal returns in all of the five differential ($Illiqid - Liquid$) portfolios are positive and statistically different from zero at the 5% level. This again supports the notion that illiquidity measured by the price-impact parameters is priced in the cross-section of stock returns.

Following Fama and French (1993), we test a hypothesis in Panel B by employing the statistic based on Gibbons, Ross, and Shanken (GRS) (1989). The null hypothesis for the GRS test is, 'H_0^{Diff}: the $FF3$ *Alphas* for the five differential ($Illiqid - Liquid$) portfolios are jointly zero.' The GRS statistic is defined as follows. Let there be M time-

[9] See Panel (A) in Table I of their paper.

series observations, G portfolios, and $L-1$ factors (excluding the intercept). Further, let X denote the matrix of regressors. Then the test statistic is given by

$$\left(A'\Sigma^{-1}A\right)\frac{M-L-G+1}{G(M-L)\omega_{1,1}},$$

where A is the column vector of FF3 Alphas, Σ is the variance-covariance matrix of the residuals from the time-series regressions, and $\omega_{1,1}$ is the diagonal element of $\left(X'X\right)^{-1}$ corresponding to the intercept. Under the null hypothesis, this statistic follows an F-distribution with G and $M-L-G+1$ degrees of freedom. As the F-statistic and the corresponding p-value suggest in Panel B, the null hypothesis is strongly rejected.

Of course, the portfolio analyses are preliminary in a sense that they do not account for other characteristics that may affect stock returns. Moreover, portfolio returns averaged across stocks may hide important aspects underlying the data at the individual level, thereby obscuring the impact of illiquidity. We address these issues in a regression framework in the next subsection.

B. Cross-Sectional Regressions

We have observed in Tables V-VI that within a given firm-size group the average of value-weighted portfolio return as well as the abnormal return is likely to increase with illiquidity, suggesting that illiquidity is a priced factor. In this section, we formally test whether our illiquidity measures have any impact on returns. As in Brennan, Chordia, and Subrahmanyam (1998), our test involves the following cross-sectional regression estimated at the monthly frequency:

$$\tilde{R}^*_{jt} = c_{0t} + \phi_t\lambda^i_{jt} + \sum_{n=1}^{N} c_{nt}Z_{njt} + \tilde{e}'_{jt},\ i = K,\ GH,\ FV,\ \text{or}\ S, \tag{20}$$

where $\tilde{R}^*_{jt}$ represents either the risk-unadjusted excess return (EXSRET) or one of the two risk-adjusted excess returns (FF3EXSRET1 and FF3EXSRET2) defined and estimated in Section II, λ^i_{jt} is one of our four illiquidity measures (λ^K, λ^{GH}, λ^{FV}, and λ^S) estimated in Section I, and Z_{njt} denotes firm characteristic n for stock j in month t.[10] As Eq.(20) indicates, we use individual stocks' price-impact parameters (as opposed to portfolio average values) in the regressions. This is to reduce data snooping biases, as Brennan, Chordia, and Subrahmanyam (1998) point out, and at the same time not to conceal possibly important information within portfolio averages (Roll, 1977). Considering that the average number of observations in the intradaily order flows used each month for each firm to estimate the monthly price-impact parameters is 3,636.5 (see Table I), the estimation errors in the price-impact parameters may not be overwhelming. To reduce the measurement error problem, however, we also use portfolio average price-impact parameters as a robustness check later in the next subsection.

Avramov and Chordia (2006) show that a constant-beta version of the Fama and French (1993) 3-factor model cannot adequately capture the predictive ability of firm characteristics in stock returns. Thus, we control for characteristics (Z_{njt}) such as book-to-market equity (BTM) and past returns (MOM1-MOM4) when we examine the impact of our illiquidity measures on FF3-adjusted returns in Eq.(20). As mentioned earlier, however, we do not use SIZE as a control variable in the regressions because it induces multicollinearity with our illiquidity measures.[11]

Now we report the standard Fama-MacBeth statistics (the time-series average of the estimated coefficients from the regression equation above and its t-statistic) in Tables VII and VIII.[12] Along with the average coefficients and t-statistics, we also list the average

[10] In Eq.(20), monthly subscript "t" is denoted as contemporaneous. Note, however, that all the explanatory variables are constructed with past data, relative to the dependant variable, which is based on the month-end price in month t. For example, λ_t^i is estimated with intradaily order flows from the begining of the first trading day to the end of the last trading day in month t, BTM is six-month lagged as defined in Section III, SIZE is log(MV), where MV is the market value of month $t-1$, and the momentum variables are basically the returns of past months.

[11] To show how SIZE causes multicollinearity with our illiquidity measures, in Table A1 we report four different regression specifications for each of the four different illiquidity measures using FF3EXSRET1 as a dependent variable. We observe in Panel A, for example, that illiquidity (measured by GRT-adjusted λ^K) is strongly positively related to stock returns (see Specification 1) while (GRT-adjusted) SIZE is strongly negatively related to returns (see Specification 2). However, when both λ^K and SIZE are included together in the regressions (see Specifications 3 and 4), the sign of the loading on λ^K reverses to negativity. There is no exception for the other three illiquidity measures (Panels B-D). Therefore, we exclude SIZE in the regression analyses all throughout the paper.

[12] Since the dependent variable in the cross-sectional regressions is the monthly stock return as in the original study of Fama and MacBeth (1973), which is close to being serially uncorrelated, we find no evidence of statistically significant autocorrelations in the time series of the estimated coefficients (the absolute values of the first-order serial correlations in the coefficient series were lower than 10%).

of the adjusted R^2 values from the individual regressions (*Avg R-sqr*) and the average number of companies used in the regression each month over the sample period (*Avg Obs*). For some nonstationary variables, we use the GRT-adjusted series (indicated by superscript "a") for the cross-sectional regressions.

The regression results with the two illiquidity measures estimated using *raw* order flows (λ^K and λ^{GH}) are presented in Table VII, while those with the other two lambdas estimated using *unexpected* order flows (λ^{FV} and λ^S) appear in Table VIII. As we see in Table VII, the average number of component stocks used in the monthly regressions for NYSE stocks ranges from 1,513.9 to 1,578.1, depending on data availability of the variables. *Avg R-sqr* is about 2.9-4.5%. The explanatory power of the regressions is higher with the unadjusted excess returns (EXSRET) than with the risk-adjusted returns (FF3EXSRET1, FF3EXSRET2). This suggests that the Fama-French model has some ability to price stocks in the cross-section. Given that FF3EXSRET2 requires more strict conditions in estimation (e.g., past return series of at least 24 months in each month), Table VII shows that the average number of component stocks is smaller in the last regression specification (FF3EXSRET2) than in the first two specifications. The patterns shown in Table VIII are similar to those in Table VII.

We first discuss the results from the Fama-MacBeth regressions of EXSRET on illiquidity estimated based on Kyle (1985), λ^K, as well as other firm characteristics that are known to be associated with returns, namely, BTM and the four momentum variables (MOM1-MOM4). The second column of Panel A in Table VII shows that the average coefficient of λ^K is positive and statistically significant at the 1% level after controlling for other firm characteristics, confirming the hypothesis that stocks with higher illiquidity provide higher (excess) returns. Consistent with the prior literature, the coefficient of BTM is positive and statistically significant. The last two momentum variables are also positively related to returns.

We now consider whether the relations observed above are maintained when the

Therefore, we report the standard Fama-MacBeth t-statistic instead of the Newey and West (NW) (1987, 1994) t-statistic throughout the paper. An unreported table equivalent to Table VII which includes both types of t-statistics shows that the levels of t-values are very similar to each other. (As suggested by NW in choosing bandwidth parameter N ($= L + 1$) for the Bartlett kernel to compute the NW standard errors, we let lag length L be equal to the integer portion of $4(T/100)^{2/9}$, where T is the number of observations in the estimated coefficient series.) The table is available upon request.

dependent variable is risk-adjusted using the FF factors. The estimates of illiquidity and characteristic rewards ($\widehat{\phi}$ and $\widehat{c}_n$) for returns adjusted by the first method in Section II (FF3EXSRET1) are presented in the next column of Panel A. By risk-adjusting, the coefficient of MOM2 becomes statistically significant. However, the key relations are essentially unchanged. λ^K continues to be strongly positively related to risk-adjusted returns. The book-to-market ratio shows a similar pattern. Overall, BTM plays an important role in predicting stock returns in the NYSE market. The coefficients of the momentum variables (except for MOM1) imply that better price performance in the past tends to provide higher returns in the current month. This finding confirms the continuation in returns documented by Jegadeesh and Titman (1993).

In the last column of Panel A, we report the estimates of illiquidity and characteristic rewards ($\widehat{\phi}$ and $\widehat{c}_n$) for excess returns (FF3EXSRET2), which are now risk-adjusted using rolling estimates of betas as described in Section II. First, the impact of λ^K on risk-adjusted returns is slightly higher than the result with FF3EXSRET1. BTM continues to have an impact on excess returns. With FF3EXSRET2, the coefficient of MOM1 turns negative, but MOM2 and MOM3 are positively related to returns.

Now we examine in Panel B the results with the illiquidity measure estimated based on the Glosten and Harris (1988) model, where trading costs are decomposed into two elements. While the roles of the firm characteristic variables are similar to those in Panel A, the impact of λ^{GH} is now much stronger. To gauge the effect of illiquidity on the stock return in the FF3EXSRET2 specification, we find that an increase in illiquidity (λ^{GH}) by one standard deviation results in higher monthly (excess) returns of 0.25%. The magnitude of the additionally required monthly returns is economically significant, given that Chordia, Huh, and Subrahmanyam (2007) document that the average monthly (raw) return is 1.19% for 1,647.2 NYSE/AMEX stocks over the past 39.5 years.[13]

Next, we investigate in Table VIII how the effects of illiquidity and other firm characteristics on returns change when we estimate the illiquidity measures based on Foster and Viswanathan (1993) and Sadka (2006), who use *unexpected* order flows. We see in

[13] If we conduct regressions using the price-impact parameters without the GRT-adjustments, the average coefficients are 85-96% larger and the t-values are 31.1-36.3% higher. In the regression specification with the dependent variable being FF3EXSRET1 as in Panel B of Table VII, for instance, when we use the parameter (λ^{GH}) without the GRT-adjustment, the average coefficient is 0.543 (vs. 0.281 with the GRT-adjustment) and the t-statistic is 7.63 (vs. 5.60 with the adjustment).

Panel A of Table VIII that the coefficients of the Foster-Viswanathan lambda, λ^{FV}, are also statistically different from zero at any conventional level after accounting for the effects of the firm characteristics. As observed before in Table VII, BTM has a similar impact on returns. The momentum effects are also similar to those in Table VII. As for the illiquidity measure based on Sadka (2006), who decomposes trading costs into *four* components and employs *unexpected* order flows, it is reasonable to see that the coefficients of λ^S and their t-values in Panel B are smallest among the last three peers. However, the Sadka lambda is also statistically significant at the 1% level. The impacts of the other variables are comparable to those reported in Panel A.[14]

C. Robustness Checks

C.1. Using Portfolio Average Price-impact Parameters

In Eq.(20), we have avoided the measurement error problem in the loadings on the FF 3 factors by using one of the two types of risk-adjusted returns as a dependent variable. But λ_{jt}^i is an explanatory variable that is estimated from the time-series regression. Given that the average number of transactions used within a month to estimate the monthly price-impact parameters is large enough (3,636.5 data points on average each month for each firm), the estimation errors in the price-impact parameters may not be substantial. Nonetheless, this error-in-variable problem could still be a concern when we use the price-impact parameters of *individual* stocks. In the spirit of Fama and French (1992), therefore, we now examine whether our empirical results are robust to using the *portfolio average* price-impact parameters.

To obtain the portfolio average price-impact parameters, each month the component stocks are first split into 10 portfolios (with the equal number of stocks) after being sorted in ascending order by firm size (MV) and then each of the 10 portfolios is again split into

[14] Spiegel and Wang (SW) (2005) document that cost-based illiquidity measures do not tend to be priced after controlling for the idiosyncratic risk. In line with their study, we have estimated the idiosyncratic risk measure (termed as *SIGMA*) against the Fama-French 3 factors using the data from the past 60 months (at least 24 months). Including SIGMA in Eq.(20), we then conduct the same type of regression analysis equivalent to Tables VII-VIII . The result shows that the coefficients of SIGMA are positive but insignificant. However, our four illiquidity measures remain significant at any conventional level even after controlling for SIGMA. Full results are available from the author.

10 portfolios after being sorted by each of the four GRT-adjusted price-impact parameters (λ^K, λ^{GH}, λ^{FV}, and λ^S), resulting in 100 portfolios for each parameter in each month. Then the portfolio average lambda is computed for each portfolio and this average value is assigned to the component stocks of that portfolio. The resulting portfolio average price-impact parameters corresponding to λ^K, λ^{GH}, λ^{FV}, and λ^S are denoted as λ_p^K, λ_p^{GH}, λ_p^{FV}, and λ_p^S, respectively.

The regression analyses with the portfolio average lambdas are contained in Table IX. For brevity, we do not report the results using the unadjusted return (EXSRET) as a dependent variable. LAM_PAVG in the table stands for one of the four portfolio average lambdas. Comparing the results in Panels A and B of Table IX with those in Table VII, we see that the size of the loadings on the two portfolio average lambdas, λ_p^K and λ_p^{GH}, increases slightly but their t-values become smaller, relative to those of λ^K and λ^{GH}. Regardless of how returns are risk-adjusted, however, both λ_p^K and λ_p^{GH} continue to be positive and statistically significant at any conventional level. The relations of the other variables to returns are essentially the same.

The comparison of Panels C and D in Table IX with the relevant panels in Table VIII presents qualitatively similar patterns for the other two measures, λ_p^{FV} and λ_p^S. In any case, the coefficients of the portfolio average parameters are statistically significant at the 1% level. As for the other variables, the differences are minimal. On balance, using the portfolio average price-impact parameters does not significantly change our key results. Based on this finding, therefore, we will continue to report the regression results using the *individual* price-impact parameters in the remaining analyses.[15]

C.2. Using Quote Mid-point Returns

A recent study by Bessembinder and Kalcheva (2006) argues that empirical pricing tests using the returns calculated based on the observed closing prices might induce microstructure biases because of the bid-ask bounce, suggesting that asset-pricing tests with quote mid-point returns can reduce this problem. To address this issue, we compute quote

[15] To obtain the portfolio average price-impact parameters, we also used book-to-market equity (BM-w) as a sorting variable, instead of firm size (MV). The results using book-to-market equity are very similar to those using firm size. The table is available upon request.

mid-point returns using the monthly closing quote mid-point prices after matching the intradaily trades and quotes via the Lee and Ready (1991) algorithm, which is described in Section I. We then risk-adjust the mid-point returns (in excess of the one-month T-bill rate) according to the two methods as before. The sample period is: 276 months (198301-200512) when the risk-adjustment in returns is performed by the first method to obtain FF3EXSRET1; and 252 months (198501-200512) when the risk-adjustment is performed by the second method to obtain FF3EXSRET2.[16]

The cross-sectional regression results with the mid-point (excess) returns are shown in Table X. To save space, the results from the specification that uses the unadjusted excess return (EXSRET) will no longer be reported throughout. LAM denotes one of the four price-impact parameters (λ^K, λ^{GH}, λ^{FV}, and λ^S). One aspect we recognize in Panels A-D is that the adjusted R^2's (Avg R-sqr) range from 0.8% to 1.1%, which are much lower than those reported in Tables VII-VIII. The other is that by using quote mid-point returns, the impacts of the explanatory variables on returns are weakened in general. For example, the book-to-market and momentum effects all phase out.[17]

An interesting feature in Table X is that while the sensitivity of the mid-point returns to the four illiquidity measures increases, the level of significance decreases substantially compared to that in Tables VII-VIII. However, the coefficients of the illiquidity measures are likely to be significant at 1-5% (except for two cases). The lower t-values might result from the elimination of the bid-ask bounce effect as well as from the narrower sample.

C.3. Using Share Order Flows

Apart from the two different tests above using the portfolio average lambdas and the quote mid-point returns, we have already employed the four different types of price-impact parameters (λ^i_{jt}, $i = K$, GH, FV, and S) as well as the three different types of excess returns (EXSRET, FF3EXSRET1, and FF3EXSRET2) to ensure the robustness

[16] Recall that the risk-adjustment using the secong method (60-month rolling regressions) requires a return series of at least (past) 24 monthly data points each month for each firm. Thus, we lose the first 24 months. Also note that for expositional convenience we use the same notations as before for the two risk-adjusted returns (FF3EXSRET1 and FF3EXSRET2), although the returns used here are different (quote mid-point returns).

[17] When firm size (SIZE) is included in the regressions, it is strongly negatively related to quote mid-point returns. However, the sign of the loadings on the illiquidity measures reverses as before.

of our results.[18] Now we briefly discuss the effects of our choices in order flows when we estimate the illiquidity measures. As pointed out in Section I, the price-impact parameters estimated with *dollar volume*-based order flows are used as the primary basis for our main analyses. Then, will our key findings reported in Tables VII-VIII change if the price-impact parameters are estimated with *share volume*-based order flows? The unreported tables show that, by replacing the measures estimated using dollar order flows with those using share order flows, the size of the t-values for the illiquidity measures decreases slightly (about 10-15%). Given the generally large levels of the t-values, however, it does not affect very much the main features observed in Tables VII-VIII. In the analyses equivalent to the two tables, we find that the coefficients of the four corresponding illiquidity measures are all positive and statistically significant at any conventional level, with the key relations of the other variables being maintained.

Thus far, we have demonstrated that the four illiquidity measures continue to be priced in the cross-section of stock returns, regardless of using: 1) portfolio average illiquidity measures; 2) quote mid-point returns; or 3) share order flows to estimate the price-impact parameters. However, two more questions still remain to be answered: (i) Do the price-impact parameters estimated using the intradaily order flows perform better than the other commonly used (il)liquidity measures in the finance literature? (ii) Do the price-impact parameters continue to be priced after accounting for the effects of other alternative (il)liquidity measures? These issues are resolved in the next section.

V. A Horse Race with Alternative Measures

A. Selection of Alternative Measures and their Relations to the Price-impact Parameters

Although it has become possible to process order flows and estimate the price-impact parameters for a broad sample of stocks as well as for a long period of time, it is still

[18]As mentioned earlier, we additionally obtained a "purged" estimator of Black et al. (1972) for each of the explanatory variables in the regressions of FF3EXSRET1 and FF3EXSRET2. The results were very similar to those of the "raw" estimator and are not reported. The results imply that the estimation errors in factor loadings are not correlated with the vector of explanatory variables. These results are available from the authors upon request.

computationally challenging. On the other hand, there are a number of other (il)liquidity measures that have been used in the asset-pricing or microstructure literature. Some of them are available with low costs and less computational burden because they can be estimated or constructed from daily or lower frequency databases. We now select some of commonly used alternative measures available with low costs and investigate how our price-impact parameters estimated from the high frequency databases compare to those alternatives.

As Hasbrouck (2005) admits, estimating the measures using high-frequency data may be limited to the relatively small and recent data samples because of data availability or computational difficulties. Merton (1980) also suggests that the accuracy in estimating first moments hinges upon the length of the data sample but not the sampling frequency. It is also relevant to recognize the computational economy of liquidity measures that can be constructed from data of daily or lower frequency. As such, given the issues described above in selecting alternative measures for comparison purposes, we limit our choices to the measures that can be estimated using the CRSP daily file.

First, we consider Amihud's (2002) illiquidity measure, which is defined as $|r|/DVOL$, where r is the daily stock return and $DVOL$ is the daily dollar volume (in \$100,000). For monthly regressions, we compute each month the average of the daily estimates of illiquidity within a month. Roughly speaking, this measure (notated as *Amihud* in our analysis) is similar to Kyle's (1985) lambda, which is one of the four price-impact parameters considered in this study. However, the Amihud measure is distinct from the Kyle lambda in the sense that *Amihud* captures the absolute return impact of *unsigned* volume, while the Kyle measure is the price impact of *signed* volume (order flows). Given the fact that the Amihud measure has been used widely in recent literature, however, we include *Amihud* as one of the competing illiquidity measures.

Attempting to answer the question of how well high-frequency measures can be proxied using daily data, Hasbrouck (2005) suggests that the market risk-adjusted effective cost of Roll (1984), estimated using the Gibbs sampler, is one of appropriate CRSP-based proxies for a TAQ-based effective cost. We thus consider this measure (notated as *Roll_Gibbs*) in our study.[19]

[19] This measure is described in Hasbrouck (2006), who denotes it as c_BMA.

Lastly, if a stock is not perfectly liquid, signed volume may induce adjustments in stock prices that initially overshoot and subsequently revert to the true values. Therefore, we estimate a reversal measure of illiquidity each month for each stock using the CRSP daily file as in Pástor and Stambaugh (2003) who estimate γ from the regression equation,

$$r^e_{j,d+1,t} = a + br_{j,d,t} + \gamma sign(r^e_{j,d,t})DVOL_{j,d,t} + \varsigma_{j,d+1,t},$$

where $r_{j,d,t}$ is the raw return and $r^e_{j,d,t}$ is the excess return (over the CRSP value-weighted index return) of stock j at day d within month t (we require at least 15 days of data per month in the CRSP daily file to estimate γ). In many cases, γ is negative, implying that stock prices often reverse on the following day. A stock with a larger absolute value ($|\gamma|$) of it is assumed to be more illiquid. We denote $|\gamma|$ as PS.

The two illiquidity measures (*Amihud* and PS), are estimated each month and the *Roll_Gibbs* measure is available from Joel Hasbrouck's website at an annual frequency. For consistency in comparisons, the three alternative measures chosen above are also Winsorized each month at the 0.5th and 99.5th percentiles before conducting the correlation and regression analyses below.

In Table XI, we report correlation between our price-impact parameters and the three alternative illiquidity measures. Our four price-impact parameters are most highly correlated with *Amihud* (37-54%), followed by PS. Especially, we find that correlation of *Roll_Gibbs* with the price-impact parameters ranges from 28.8% to 41.1%. This demonstrates that, contrary to the argument by Hasbrouck (2005), *Roll_Gibbs* is at best a noisy proxy for a TAQ-based effective cost, justifying the merits of estimating the price-impact parameters using high-frequency databases such as the TAQ. The table also shows that *Amihud* is more highly correlated with PS and *Roll_Gibbs* than with our price-impact parameters.

B. *Cross-Sectional Regressions with Alternative Illiquidity Measures*

In this subsection, we conduct a horse race between one of our four price-impact parameters and one (or all) of the three alternative measures considered in the previous subsection. Our goal is to test whether the effects of our illiquidity measures on returns are comparable to those of the other alternative measures and, going one step further, to check whether each of the price-impact parameters still has an incremental impact on returns after accounting for the effects of the three alternative measures.

First, we run the regressions with each of the three alternative measures by replacing λ^i_{jt} in Eq.(20) with one of *Amihud*, *Roll_Gibbs*, and *PS*. Since *Roll_Gibbs* is at an annual frequency, we keep the annual values of this measure constant over the twelve months within each year for the monthly regressions. We report the results in Table XII. *ALT* in the table stands for one of the three alternative measures.[20]

As the correlation coefficients in Table XI suggest, Panel A of Table XII shows that *Amihud* is strongly positively related to returns. The impact of *Amihud* on returns is comparable to that of our four price-impact parameters reported in Tables VII-VIII. However, the impact of *Roll_Gibbs* in Panel B has the wrong sign (negative) and statistically insignificant. This is surprising, given that this measure is highly correlated with *Amihud*. As Hasbrouck (2006) indicates, the limitation of *Roll_Gibbs* stems from the fact that it does not explicitly incorporate the price-impact effects of trading volume or order flows, which may be endogenous with price dynamics. Moreover, the illiquidity effect of asymmetric information is not likely to be captured by the bid-ask spread, but by the price-impact of a trade. The negative sign on *Roll_Gibbs* is consistent with Eleswarapu and Reinganum (1993) and Chordia, Huh, and Subrahmanyam (2008). Panel C shows that the impact of *PS* on returns is positive but marginal.[21]

Next, we run a horse race between one of our four price-impact parameters and all the three competing measures together. For this purpose, we augment Eq.(20) by including three more variables as in the following equation,

$$\tilde{R}^*_{jt} = c_{0t} + \phi_t \lambda^i_{jt} + \sum_{s=1}^{3} \varphi_{st} ALT_{sjt} + \sum_{n=1}^{N} c_{nt} Z_{njt} + \tilde{e}'_{jt},\ i = K,\ GH,\ FV, \text{ or } S, \tag{21}$$

where ALT_{sjt} ($s = 1$, 2, and 3) denotes one of the three alternative illiquidity measures

[20] Note that the three alternative measures are GRT-adjusted for each firm over the sample period before conducting the cross-sectional regressions in order to eliminate nonstationarity.

[21] This result is consistent with Hasbrouck (2005) (see his Table 5). Unlike this study, however, Pástor and Stambaugh (2003) use their measure to price illiquidity risk, rather than the level of illiquidity itself.

($Amihud$, $Roll\_Gibbs$, and PS).

As shown in Panel A of Table XIII, by including the additional illiquidity measures in the regressions, the number of average component stocks decreases but the adjusted R^2 increases, compared to the results in Panel A of Table VII. We also observe in Panel A that λ^K remains to be priced even after controlling for the three alternative measures. The effect of $Amihud$ is similar to that of Panel A in Table XII, although there are some differences in the size of the coefficients or t-values. The effect of $Roll\_Gibbs$ is negative and marginally significant, but PS plays no role. With the additional illiquidity measures, the book-to-market effect is maintained. While MOM1 is negatively related to returns, the impacts of the other momentum variables become weakened. Panels B-D of Table XIII report the analogs of Panel A using the other three illiquidity measures (λ^{GH}, λ^{FV}, and λ^S). We find that the effects of these three price-impact parameters continue to be positive and statistically significant at the 1% level. The other variables show the patterns similar to those in Panel A.

To sum up, our empirical tests provide solid evidence that illiquidity measured by the four price-impact parameters is a priced attribute in the cross-section of stock returns, even after controlling for the alternative measures. In addition, while estimating the price-impact measures is most onerous, the unambiguously consistent relations of the parameters to returns, together with the strong levels of their statistical significance throughout a variety of experiments, indicate that the price-impact parameters are among the most reliable proxies for illiquidity in stock markets.

VI. Conclusion

A number of empirical proxies for (il)liquidity have been proposed in the literature, which links trading frictions to asset prices. However, these measures have been subject to controversy because they have achieved equivocal results when answering the question of whether (il)liquidity is related to asset returns. Moreover, many of the measures have often been constructed from low frequency databases, which may conceal some interesting dynamics in the aggregation procedures. In cases where illiquidity measures are estimated from high frequency databases, the coverage has not been broad or long enough because of limited data availability or computational burden. As a result, some theory-based illiquidity measures have been derived and tested. But there is a debate about the appropriateness of some empirical proxies used as inputs to these theoretical illiquidity measures.

With these issues in mind, we choose rather a direct and intuitive approach to measuring illiquidity. Specifically, we estimate a set of price-impact parameters based on the four different models using the intradaily order flows. The order flows are processed through the Lee and Ready (1991) algorithm from the ISSM and TAQ databases. The coverage is comprehensive and long enough, spanning the past 23 years (276 months) for more than 1,500 NYSE-listed stocks.

Our empirical analyses lend strong support to the notion that illiquidity measured by the price-impact parameters is priced in the cross-section of returns, even after accounting for the effects of risk factors, firm characteristics, and other illiquidity proxies prevalent in the literature. Given the consistent relations of our illiquidity measures to returns, the results also suggest that the price-impact parameters estimated using the high-frequency order flow data are more effective proxies for illiquidity, compared with other alternative measures constructed from data of daily or lower frequency.

References

Acharya, V., and L. Pedersen, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 385-410.

Admati, A. and P. Pfleiderer, 1988, A theory of intraday patterns: Volume and price variability, *Review of Financial Studies* 1, 3-40.

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

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

Arbel, A., and P. Strebel, 1982, The neglected and small firm effects, *Financial Review* 17, 201-218

Atkins, A., and E. Dyl, 1997, Market structure and reported trading volume: NASDAQ versus the NYSE, *Journal of Financial Research* 20, 291-304.

Avramov, D. and T. Chordia, 2006, Asset pricing models and financial market anomalies, *Review of Financial Studies* 19, 1002-1040.

Bessembinder, H., and I. Kalcheva, 2006, Liquidity biases in asset pricing tests, working paper, University of Utah.

Black, F., Jensen, M., and M. Scholes, 1972, The capital asset pricing model: Some empirical tests, In: Jensen, M. (Ed.), *Studies in the Theory of Capital Markets*, Praeger Publishers, New York.

Brennan, M., Chordia, T., and A. Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected stock returns, cross-sectional determinants of expected returns, *Journal of Financial Economics* 49, 345–373.

Brennan, M., and A. Subrahmanyam, 1996, Market microstructure and asset pricing: On the compensation for illiquidity in stock returns, *Journal of Financial Economics* 41, 441-464.

Chordia, T., S. Huh, and A. Subrahmanyam, 2007, The cross-section of expected trading activity, *Review of Financial Studies* 20, 709-740.

Chordia, T., S. Huh, and A. Subrahmanyam, 2008, Theory-based illiquidity and asset pricing, *Review of Financial Studies,* Forthcoming.

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

Chordia, T., R. Roll, and A. Subrahmanyam, 2002, Order imbalance, liquidity, and market returns, *Journal of Financial Economics* 65, 111-130.

Connor, G., 1984, A unified beta pricing theory, *Journal of Economic Theory* 34, 13-31.

Easley, D., and M. O'Hara, 1987, Price, trade size, and information in securities markets, *Journal of Financial Economics* 19, 69-90.

Eisfeldt, A., 2004, Endogenous liquidity in asset markets, *Journal of Finance* 59, 1-30.

Eleswarapu, V. R., and M. Reinganum, 1993, The seasonal behavior of the liquidity premium in asset pricing, *Journal of Financial Economics* 34, 373-386.

Fama, E., and K. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-466.

Fama, E., and K. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, E., and J. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.

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

Gallant, R., P. Rossi, and G. Tauchen, 1992, Stock prices and volume, *Review of Financial Studies* 5, 199-242.

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

Gibbons, M., S. Ross, and J. Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121-1152.

Glosten, L., 1989, Insider trading, liquidity, and the role of the monopolist specialist, *Journal* *of Business* 62, 211-235.

Glosten, L., and L. Harris, 1988, Estimating the components of the bid-ask spread, *Journal of Financial Economics* 21, 123-142.

Hasbrouck, J., 1999, The dynamics of discrete bid and ask quotes, *Journal of Finance* 54, 2109-2142.

Hasbrouck, J., 2005, Trading costs and returns for US equities: The evidence from daily data, working paper, New York University.

Hasbrouck, J., 2006, Trading costs and returns for US equities: Estimating effective costs from daily data, working paper, New York University.

Jacoby, G., D. Fowler, and A. Gottesman, 2000, The capital asset pricing model and the liquidity effect: A theoretical approach, *Journal of Financial Markets* 3, 69-81.

Jegadeesh, N., and S. Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48,65-92.

Johnson, T., 2005, Dynamic liquidity in endowment economies, *Journal of Financial Economics*, 80, 531-562.

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

Kyle, A., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315-1335.

Kumar, P., and D. Seppi, 1994, Information and index arbitrage, *Journal of Business* 67, 481-509.

Lee, C., 1993, Market fragmentation and price-execution in NYSE-listed securities, *Journal of Finance* 48, 1009-1038.

Lee, C., and M. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance* 46, 733-747.

Lee, C., and B. Radhakrishna, 2000, Inferring investor behavior: evidence from TORQ data, *Journal of Financial Markets* 3, 83-111.

Lo, A., and C. MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431-468.

Mayshar, J., 1983, On divergence of opinion and imperfections in capital markets, *American Economic Review* 73, 114-128.

Merton, R., 1980, On estimating the expected rate of return on the market, *Journal of Financial Economics* 8, 323-362.

Merton, R., 1987, A simple model of capital market equilibrium with incomplete information, *Journal of Finance* 42, 483-510.

Miller, E., 1977, Risk, uncertainty, and divergence of opinion, *Journal of Finance* 32, 1151-1168.

Newey, W., and K. West, 1987, A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.

Newey, W., and K. West, 1994, Automatic lag selection in covariance matrix estimation, *Review of Economic Studies* 61, 631-653.

Odders-White, E., 2000, On the occurrence and consequences of inaccurate trade classification, *Journal of Financial Markets* 3, 205-332.

Pástor, L., and R. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 113, 642-685.

Roll, R., 1977, A critique of the asset pricing theory's tests: on past and potential testability of theory, *Journal of Financial Economics* 4, 129-176.

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

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

Shanken, J., 1985, Multibeta CAPM or equilibrium APT? A reply, *Journal of Finance* 40, 1189-1196.

Shanken, J., 1987, Multivariate proxies and arbitrage pricing relations: living with the Roll critique, *Journal of Financial Economics* 18, 91-110.

Spiegel, M., and X. Wang, 2005, Cross-sectional variation in stock returns: liquidity and idiosyncratic risk, working paper, Yale University.

Table I

Summary of Intra-Daily Order Flow Data Used to Estimate the Price-Impact Parameters for NYSE-Listed Stocks

This table reports the summary of intra-daily order flow data used to estimate price-impact parameters (or *lambdas*) for NYSE-listed stocks. To obtain the order flow data, each trade is classified as a buyer- or seller-initiated trade according to the Lee and Ready (1991) algorithm using trades and quotes data from the Institute for the Study of Securities Markets (ISSM: 1983-1992) and the NYSE Trades and Automated Quotations (TAQ: 1993-2005) databases. In order to match trades and quotes, any quote less than five seconds prior to the trade is ignored and the first one at least five seconds prior to the trade is retained (a *five*-second rule) for the years 1983 to 1998. For the last seven years (1999-2005), a *two*-second rule is applied. Then the trades are signed as follows. If a trade occurs above the prevailing quote mid-point, it is classified as buyer-initiated and *vice versa*. If a trade occurs exactly at the quote mid-point, the trade is discarded (hence such a trade is not counted in this table). Only NYSE stocks are included in the sample. To survive in the sample, stocks should have at least 110 trades per month (on average 5 trades per day). The sample period is the past 276 months (23 years: 198301-200512).

Summary of Intra-Daily Order Flow Data Used to Estimate Price-Impact Parameters	
Total Number of Trades (and Matched Quotes) Used over the Sample Period (198301-200512):	2,000,884,544
Minimum Number of Monthly Trades (and Matched Quotes) per Firm:	110
Maximum Number of Monthly Trades (and Matched Quotes) per Firm:	173,471*
Average Number of Monthly Trades (and Matched Quotes) per Firm:	3,636.5
Pooled Number of Firms Used over the 276 Months (Firm-month Observations):	443,453

*The number of trades (per month) in shares of Chevron Corporation in October 2005.

Table II

Sign, Statistical Significance, and Descriptive Statistics of the Price-Impact Parameters Estimated for NYSE-Listed Stocks

This table reports the sign and statistical significance of the estimated price-impact parameters (Panel A) and the descriptive statistics (average Mean, Median, Standard Deviation (STD), Coefficient of Variation (CV), Skewness, and Kurtosis) for the parameters (Panel B). The price-impact parameters are estimated based on the four different models using the intradaily order flow data processed via the Lee and Ready (1991) algorithm. Each price-impact parameter is defined as follows: λ_0^K : the price-impact parameter estimated based on Kyle (1985) (multiplied by 10^6) using intradaily dollar order flows available within each month for each stock; λ^K : the Winsorized parameter of λ_0^K at the 0.5th and 99.5th percentiles; λ_0^{GH} : the price-impact parameter estimated based on Glosten and Harris (1988) (multiplied by 10^6) using intradaily dollar order flows available within each month for each stock; λ^{GH} : the Winsorized parameter of λ_0^{GH} at the 0.5th and 99.5th percentiles; λ_0^{FV} : the price-impact parameter estimated based on Foster and Viswanathan (1993) (multiplied by 10^6) using intradaily dollar order flows available within each month for each stock; λ^{FV} : the Winsorized parameter of λ_0^{FV} at the 0.5th and 99.5th percentiles; λ_0^S : the price-impact parameter estimated based on Sadka (2006) (multiplied by 10^6) using intradaily dollar order flows available within each month for each stock; λ^S : the Winsorized parameter of λ_0^S at the 0.5th and 99.5th percentiles. To survive in the sample, stocks should have at least 110 trades per month (on average 5 trades per day). In Panel B, the mean values of each statistic are first calculated cross-sectionally each month and then the time-series averages of those values are reported here. The sample period is the past 276 months (23 years: 198301-200512) for NYSE stocks. The average number of component stocks used each month is 1,578.1.

Panel A: Sign and Significance of the Estimated Price-Impact Parameters

Estimation Method	Notation	% of Positive Lambdas	% of Positive & Significant Lambdas (at 5%)	% of Negative Lambdas
Kyle (1985)	λ_0^K	99.78%	97.19%	0.22%
Glosten and Harris (1988)	λ_0^{GH}	97.97%	85.24%	2.03%
Foster and Viswanathan (1993)	λ_0^{FV}	97.68%	83.99%	2.32%
Sadka (2006)	λ_0^S	93.37%	67.17%	6.63%

Panel B: Descriptive Statistics for the Price-Impact Parameters

Price-Impact Parameter	Mean	Median	STD	CV	Skewness	Kurtosis
λ_0^K	1.027	0.254	7.359	456.51	14.46	432.01
λ^K	0.817	0.254	1.425	177.24	3.43	14.68
λ_0^{GH}	0.599	0.111	6.206	677.39	15.46	496.09
λ^{GH}	0.417	0.111	0.843	202.56	3.99	19.86
λ_0^{FV}	0.609	0.108	6.829	726.89	15.76	513.22
λ^{FV}	0.413	0.108	0.844	204.65	4.02	20.08
λ_0^S	0.288	0.082	3.211	1277.29	3.11	357.86
λ^S	0.305	0.082	0.640	210.88	3.72	17.60

Table III
Descriptive Statistics of Other Key Variables

This table reports descriptive statistics (Mean, Median, Standard Deviation (STD), Coefficient of Variation (CV), Skewness, and Kurtosis) for the key characteristic-related variables to be used in the Fama-MacBeth (1973) cross-sectional regressions. Each variable is defined as follows: *MV*: the market value defined as the previous month-end stock price times the number of shares outstanding (in $million); *SIZE*: natural logarithm of MV; *BM_w*: the Winsorized value (at the 0.5th and 99.5th percentiles) of a book-to-market ratio (BM = BV/MV), where the book value (BV) is defined as common equity plus deferred taxes in $million; *BTM*: natural logarithm of BM_w; *MOM1*: compounded holding period return of a stock over the most recent 3 months (from month *t-1* to month *t-3*); *MOM2*: compounded holding period return over the next recent 3 months (from month *t-4* to month *t-6*); *MOM3*: compounded holding period return over the 3 months from month *t-7* to month *t-9*; *MOM4*: compounded holding period return over the 3 months from month *t-9* to month *t-12*. The sample period is the past 276 months (23 years: 198301-200512) for NYSE stocks. The values of each statistic are first calculated cross-sectionally each month and then the time-series averages of those values are reported here. The average number of component stocks used in a month to compute the statistics for each variable is 1,578.1.

Variables	Mean	Median	STD	CV	Skewness	Kurtosis
MV	3189.32	677.30	10259.28	302.46	9.30	127.92
SIZE	6.44	6.43	1.71	26.64	0.07	-0.12
BM_w	0.74	0.64	0.55	73.47	2.69	14.99
BTM	-0.55	-0.46	0.73	-163.27	-0.85	2.25
MOM1	0.022	0.028	0.191	-1679.40	-0.49	8.23
MOM2	0.023	0.029	0.185	665.27	-0.37	6.99
MOM3	0.024	0.029	0.182	425.98	-0.29	6.36
MOM4	0.025	0.030	0.181	-240.27	-0.20	6.00

Table IV
Correlations between Key Variables

This table reports the average correlations between the key explanatory variables for NYSE stocks over the 276 months (23 years: 198301-200512). The cross-sectional correlation coefficients are first calculated each month and then the time-series averages of those values over the sample periods are reported here. The definitions of the variables are as follows: λ^{K} : the price-impact parameter estimated based on Kyle (1985) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5th and 99.5th percentiles; λ^{GH} : the price-impact parameter estimated based on Glosten and Harris (1988) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5th and 99.5th percentiles; λ^{FV} : the price-impact parameter estimated based on Foster and Viswanathan (1993) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5th and 99.5th percentiles; λ^{S} : the price-impact parameter estimated based on Sadka (2006) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5th and 99.5th percentiles; *SIZE*: natural logarithm of MV, which is the market value defined as the previous month-end stock price times the number of shares outstanding (in $million); *BTM*: natural logarithm of BM_w, which is the Winsorized value (at the 0.5th and 99.5th percentiles) of a book-to-market ratio (BM = BV/MV), where the book value (BV) is defined as common equity plus deferred taxes in $million; *MOM1*: compounded holding period return of a stock over the most recent 3 months (from month *t-1* to month *t-3*); *MOM2*: compounded holding period return over the next recent 3 months (from month *t-4* to month *t-6*); *MOM3*: compounded holding period return over the 3 months from month *t-7* to month *t-9*; *MOM4*: compounded holding period return over the 3 months from month *t-9* to month *t-12*. The values of each statistic are first calculated cross-sectionally each month and then the time-series averages of those values are reported here. The average number of component stocks used in a month to compute the statistics for each variable is 1,578.1.

	λ^{K}	λ^{GH}	λ^{FV}	λ^{S}	SIZE	BTM	MOM1	MOM2	MOM3	MOM4
λ^{K}	1									
λ^{GH}	0.936	1								
λ^{FV}	0.931	0.996	1							
λ^{S}	0.771	0.802	0.803	1						
SIZE	-0.569	-0.499	-0.494	-0.451	1					
BTM	0.043	0.030	0.030	0.030	-0.156	1				
MOM1	-0.051	-0.017	-0.017	-0.036	0.104	0.051	1			
MOM2	-0.063	-0.028	-0.028	-0.047	0.106	0.043	0.038	1		
MOM3	-0.061	-0.030	-0.030	-0.043	0.103	0.012	0.050	0.029	1	
MOM4	-0.058	-0.029	-0.029	-0.037	0.094	-0.028	0.041	0.043	0.025	1

Table V
Average of Firm Size, Price-Impact Parameters, and Value-Weighted Monthly Returns of the 25 Portfolios Formed by Sorting on Illiquidity and Firm Size

This table reports the average of firm size, price-impact parameters, and value-weighted monthly returns for the 25 portfolios formed by sorting on illiquidity and firm size. The component stocks are first split into 5 portfolios (with the equal number of stocks) after being sorted in ascending order by illiquidity (λ^{GH}) and then each of the 5 portfolios is again split into 5 portfolios after being sorted by firm size (MV), resulting in 25 portfolios each month. λ^{GH} is the price-impact parameter estimated based on Glosten and Harris (1988) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5^{th} and 99.5^{th} percentiles. MV is the market value (firm size) defined as the previous month-end stock price times the number of shares outstanding (in $million). Panel A contains the average of firm size (MV) (the time-series average of the cross-sectional means), while Panel B does the average of price-impact parameters for each portfolio. The upper part of Panel C contains the time-series average of value-weighted cross-sectional mean returns (monthly) for each portfolio, together with their *t*-statistics (italicized). The lower part of Panel C shows the time-series average returns for the five differential (*Illiquid - Liquid*) portfolios [each of which contains a time-series of return differences between the highest illiquidity portfolio (*Illiquid*) and the lowest illiquidity portfolio (*Liquid*) within a given firm-size group], together with their *t*-statistics to test the null hypothesis that the time-series average of the return differences equals zero. The sample period is the past 276 months (23 years: 198301-200512) for NYSE stocks. The average number of component stocks in each portfolio in a month is 63.41.

Average of Firm Size, Price-Impact Parameters, and Value-Weighted Returns					
	Firm-Size (MV) Group				
Illiquidity (λ^{GH}) Group	1 Small	2	3	4	5 Big
	Panel A: Average of Firm Size (MV)				
1 Liquid	388.24	1520.81	3526.89	7959.54	37269.18
2	349.83	968.98	1814.32	3331.97	10223.15
3	185.54	470.06	816.03	1373.93	4089.80
4	88.03	211.32	367.16	621.93	1854.58
5 Illiquid	30.43	74.77	129.54	236.29	859.11
	Panel B: Average of Price-Impact Parameters (λ^{GH})				
1 Liquid	-0.011	0.008	0.012	0.013	0.012
2	0.050	0.049	0.048	0.046	0.042
3	0.122	0.120	0.116	0.111	0.104
4	0.345	0.328	0.314	0.300	0.276
5 Illiquid	2.475	1.743	1.421	1.239	1.090
	Panel C: Average of Value-Weighted Returns				
1 Liquid	0.0142	0.0138	0.0153	0.0147	0.0162
	4.49	*4.95*	*5.96*	*6.02*	*6.57*
2	0.0097	0.0131	0.0167	0.0176	0.0197
	2.86	*4.40*	*5.53*	*6.33*	*7.39*
3	0.0094	0.0135	0.0156	0.0172	0.0205
	2.50	*4.13*	*5.29*	*6.02*	*7.38*
4	0.0096	0.0145	0.0166	0.0189	0.0198
	2.42	*4.17*	*5.19*	*6.72*	*6.84*
5 Illiquid	0.0241	0.0198	0.0199	0.0202	0.0207
	4.55	*5.08*	*5.86*	*6.87*	*7.93*
Illiquid - Liquid	0.0098	0.0060	0.0046	0.0055	0.0046
	2.48	*2.12*	*1.90*	*2.70*	*2.14*

Table VI

Intercepts (Abnormal Returns, *FF3 Alphas*) from the Time-series Regressions of Returns on the Fama-French 3 Factors for the 25 Portfolios Formed by Sorting on Illiquidity and Firm Size

This table reports the intercepts (abnormal returns, termed as *FF3 Alphas*), together with the *t*-statistics (italicized), for the 25 portfolios formed by sorting on illiquidity and firm size. The component stocks are first split into 5 portfolios (with the equal number of stocks) after being sorted in ascending order by illiquidity (λ^{GH}) and then each of the 5 portfolios is again split into 5 portfolios after being sorted by firm size (MV), resulting in 25 portfolios each month. Then the cross-sectional mean return (value-weighted) is calculated each month for each portfolio. λ^{GH} is the price-impact parameter estimated based on Glosten and Harris (1988) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5^{th} and 99.5^{th} percentiles. MV is the market value (firm size) defined as the previous month-end stock price times the number of shares outstanding (in $million). The 25 *FF3 Alphas* contained in the upper part of Panel A are the intercepts from the time-series regressions of the portfolio return (in excess of the one-month T-bill rate) on the Fama-French 3 factors. The values italicized in the second row of each illiquidity group in the upper part of Panel A are *t*-statistics from the time-series regressions. The five *FF3 Alphas* contained in the lower part of Panel A are the intercepts from the time-series regressions of the return difference [between the highest illiquidity portfolio (*Illiquid*) and the lowest illiquidity portfolio (*Liquid*) within a given firm-size group] on the Fama-French 3 factors, together with their *t*-statistics from the regressions. Panel B reports the GRS-test result for the null hypothesis, ‘ H_0^{Diff} : the *FF3 Alphas* for the five (*Illiquid - Liquid*) portfolios are jointly zero.’ The *F*-statistic is computed based on Gibbons, Ross, and Shanken (1989) and the corresponding *p*-value is also reported for statistical inference. The sample period is the past 276 months (23 years: 198301-200512) for NYSE stocks. The average number of component stocks in each portfolio in a month is 63.41.

Panel A: FF3 Alphas from the Time Series of Value-Weighted Returns

Illiquidity (λ^{GH}) Group	Firm-Size (MV) Group: 1 Small	2	3	4	5 Big
1 Liquid	-0.0003	-0.0002	0.0023	0.0026	0.0061
	-0.18	*-0.11*	*1.98*	*2.79*	*9.58*
2	-0.0059	-0.0010	0.0025	0.0046	0.0086
	-3.88	*-0.73*	*1.85*	*4.62*	*8.87*
3	-0.0065	-0.0009	0.0022	0.0037	0.0087
	-3.32	*-0.57*	*1.80*	*2.89*	*7.90*
4	-0.0060	-0.0006	0.0023	0.0059	0.0074
	-2.85	*-0.40*	*1.61*	*4.51*	*5.04*
5 Illiquid	0.0090	0.0053	0.0065	0.0071	0.0097
	2.43	*2.60*	*3.69*	*5.15*	*6.51*
Illiquid - Liquid	0.0093	0.0055	0.0042	0.0045	0.0036
	2.46	*2.32*	*2.25*	*3.15*	*2.34*

Panel B: GRS Test for Five (Illiquid - Liquid) Portfolios

Null Hypothesis	*F*-statistic	*p*-value
H_0^{Diff} : FF3 Alphas for five (Illiquid - Liquid) portfolios are jointly zero	*4.25*	*0.0009*

Table VII

Results of Monthly Cross-sectional Regressions: with Price-Impact Parameters Based on Kyle (1985, λ^K) and Glosten and Harris (1988, λ^{GH})

This table reports the monthly Fama-MacBeth (1973)-type cross-sectional regressions using λ^K (in Panel A) and λ^{GH} (in Panel B) for NYSE stocks over the 276 months (23 years: 198301-200512).

The dependent variable is EXSRET, FF3EXSRET1, or FF3EXSRET2 in each panel. The definitions of the variables are as follows: *EXSRET*: the monthly risk-unadjusted excess return, i.e., the monthly return less the risk-free rate proxied by the one-month T-bill rate; *FF3EXSRET1*: the risk-adjusted excess return using the Fama-French (FF) 3 factors, i.e., the constant term plus the residual from the time-series regression of the excess return on the FF 3 factors using the *entire* sample range of the data; *FF3EXSRET2*: the risk-adjusted excess return using the Fama-French (FF) 3 factors with factor loadings being estimated from the 5-year rolling regressions, i.e., R_i^* computed each month with the current month data from the equation, $R_i^* = (R_i - R_f) - [\hat{\beta}_1 MKT + \hat{\beta}_2 SMB + \hat{\beta}_3 HML]$, after the factor loadings ($\alpha, \beta_1, \beta_2, \beta_3$) are first estimated for *each month* using the time-series data of the past 60 months (at least 24 months) in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \varepsilon$, where R_i, R_f, and R_m are the individual stock return, the risk-free rate, and the market index return, respectively, while *MKT*, *SMB*, and HML are FF 3 factors; λ^K: the price-impact parameter estimated based on Kyle (1985) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5th and 99.5th percentiles; λ^{GH}: the price-impact parameter estimated based on Glosten and Harris (1988) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5th and 99.5th percentiles; *BTM*: natural logarithm of BM_w, which is the Winsorized value (at the 0.5th and 99.5th percentiles) of a book-to-market ratio (BM = BV/MV), where the book value (BV) is defined as common equity plus deferred taxes in $million and the market value (MV) is defined as the previous month-end stock price times the number of shares outstanding (in $million); *MOM1*: compounded holding period return of a stock over the most recent 3 months (from month *t-1* to month *t-3*); *MOM2*: compounded holding period return over the next recent 3 months (from month *t-4* to month *t-6*); *MOM3*: compounded holding period return over the 3 months from month *t-7* to month *t-9*; *MOM4*: compounded holding period return over the 3 months from month *t-9* to month *t-12*. To remove nonstationarity, the Gallant, Rossi, and Tauchen (GRT) (1993) procedure has been applied to λ^K, λ^{GH}, and BTM for each firm over the sample period before conducting the cross-sectional regressions (GRT-adjusted variables are indicated by superscript "a"). The values in the first row for each explanatory variable are the time-series averages of coefficients obtained from the month-by-month cross-sectional regressions, and the values italicized in the second row of each variable are *t*-statistics computed based on Fama-MacBeth (1973). The coefficients are all multiplied by 100. *Avg R-sqr* is the average of adjusted R-squared. *Avg Obs* is the monthly average number of companies used in the cross-sectional regressions. The average number of component stocks used each month in the regressions is 1,513.9-1,578.1. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

(Table VII continued)

Panel A: with Lambdas Based on Kyle (1985)							Panel B: with Lambdas Based on Glosten and Harris (1988)						
Expla. Variables	EXSRET		FF3EXSRET1		FF3EXSRET2		Expla. Variables	EXSRET		FF3EXSRET1		FF3EXSRET2	
Intercept	1.014	**	0.290	**	0.341	**	Intercept	1.019	**	0.292	**	0.349	**
	4.16		*3.01*		*3.18*			*4.18*		*3.07*		*3.30*	
$\lambda^{K,a}$	0.143	**	0.149	**	0.159	**	$\lambda^{GH,a}$	0.266	**	0.281	**	0.298	**
	3.52		*4.34*		*4.40*			*4.42*		*5.60*		*5.50*	
BTM^a	0.392	**	0.319	**	0.306	**	BTM^a	0.396	**	0.320	**	0.310	**
	6.54		*7.27*		*6.47*			*6.58*		*7.29*		*6.54*	
MOM1	-0.493		-0.578		-0.838	*	MOM1	-0.528		-0.614		-0.878	*
	-1.18		*-1.73*		*-2.23*			*-1.25*		*-1.83*		*-2.32*	
MOM2	0.562		0.738	*	0.849	**	MOM2	0.534		0.708	*	0.817	*
	1.57		*2.56*		*2.58*			*1.48*		*2.44*		*2.47*	
MOM3	0.749	*	0.808	*	0.839	*	MOM3	0.737	*	0.799	*	0.831	*
	2.04		*2.56*		*2.37*			*1.99*		*2.53*		*2.33*	
MOM4	0.994	**	0.939	**	0.588		MOM4	0.978	**	0.914	**	0.561	
	3.01		*3.52*		*1.88*			*2.93*		*3.39*		*1.78*	
Avg R-sqr	0.045		0.030		0.033		Avg R-sqr	0.044		0.029		0.032	
Avg Obs	1577.9		1576.7		1513.8		Avg Obs	1578.1		1577.0		1513.9	

Table VIII

Results of Monthly Cross-sectional Regressions: with Price-Impact Parameters Based on Foster and Viswanathan (1993, λ^{FV}) and Sadka (2006, λ^{S})

This table reports the monthly Fama-MacBeth (1973)-type cross-sectional regressions using λ^{FV} (in Panel A) and λ^{S} (in Panel B) for NYSE stocks over the 276 months (23 years: 198301-200512). The dependent variable is EXSRET, FF3EXSRET1 or FF3EXSRET2 in each panel. The definitions of the variables are as follows: *EXSRET*: the monthly risk-unadjusted excess return, i.e., the monthly return less the risk-free rate proxied by the one-month T-bill rate; *FF3EXSRET1*: the risk-adjusted excess return using the Fama-French (FF) 3 factors, i.e., the constant term plus the residual from the time-series regression of the excess return on the FF 3 factors using the *entire* sample range of the data; *FF3EXSRET2*: the risk-adjusted excess return using the Fama-French (FF) 3 factors with factor loadings being estimated from the 5-year rolling regressions, i.e., R_i^* computed each month with the current month data from the equation, $R_i^* = (R_i - R_f) - [\hat{\beta}_1 MKT + \hat{\beta}_2 SMB + \hat{\beta}_3 HML]$, after the factor loadings ($\alpha, \beta_1, \beta_2, \beta_3$) are first estimated for *each month* using the time-series data of the past 60 months (at least 24 months) in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \varepsilon$, where R_i, R_f, and R_m are the individual stock return, the risk-free rate, and the market index return, respectively, while *MKT*, *SMB*, and HML are FF 3 factors; λ^{FV} : the price-impact parameter estimated based on Foster and Viswanathan (1993) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5th and 99.5th percentiles; λ^{S} : the price-impact parameter estimated based on Sadka (2006) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5th and 99.5th percentiles; *BTM*: natural logarithm of BM_w, which is the Winsorized value (at the 0.5th and 99.5th percentiles) of a book-to-market ratio (BM = BV/MV), where the book value (BV) is defined as common equity plus deferred taxes in $million and the market value (MV) is defined as the previous month-end stock price times the number of shares outstanding (in $million); *MOM1*: compounded holding period return of a stock over the most recent 3 months (from month *t-1* to month *t-3*); *MOM2*: compounded holding period return over the next recent 3 months (from month *t-4* to month *t-6*); *MOM3*: compounded holding period return over the 3 months from month *t-7* to month *t-9*; *MOM4*: compounded holding period return over the 3 months from month *t-9* to month *t-12*. To remove nonstationarity, the Gallant, Rossi, and Tauchen (GRT) (1993) procedure has been applied to λ^{FV} , λ^{S} , and BTM for each firm over the sample period before conducting the cross-sectional regressions (GRT-adjusted variables are indicated by superscript "a"). The values in the first row for each explanatory variable are the time-series averages of coefficients obtained from the month-by-month cross-sectional regressions, and the values italicized in the second row of each variable are *t*-statistics computed based on Fama-MacBeth (1973). The coefficients are all multiplied by 100. *Avg R-sqr* is the average of adjusted R-squared. *Avg Obs* is the monthly average number of companies used in the cross-sectional regressions. The average number of component stocks used each month in the regressions is 1,513.3-1,577.4. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

(Table VIII continued)

Panel A: with Lambdas Based on Foster and Viswanathan (1993)						
Expla. Variables	EXSRET		FF3EXSRET1		FF3EXSRET2	
Intercept	1.018	**	0.292	**	0.351	**
	4.18		*3.07*		*3.32*	
λ^{FV}, a	0.270	**	0.284	**	0.298	**
	4.21		*5.28*		*5.12*	
BTM[a]	0.398	**	0.322	**	0.312	**
	6.61		*7.35*		*6.60*	
MOM1	-0.522		-0.605		-0.868	*
	-1.24		*-1.80*		*-2.29*	
MOM2	0.546		0.720	*	0.822	*
	1.51		*2.48*		*2.48*	
MOM3	0.750	*	0.810	*	0.843	*
	2.03		*2.56*		*2.37*	
MOM4	1.004	**	0.937	**	0.586	
	3.03		*3.50*		*1.86*	
Avg R-sqr	0.044		0.029		0.032	
Avg Obs	1577.4		1576.3		1513.3	

Panel B: with Lambdas Based on Sadka (2006)						
Expla. Variables	EXSRET		FF3EXSRET1		FF3EXSRET2	
Intercept	1.057	**	0.325	**	0.384	**
	4.34		*3.42*		*3.60*	
λ^{S}, a	0.203	**	0.241	**	0.257	**
	2.64		*3.74*		*3.68*	
BTM[a]	0.399	**	0.323	**	0.312	**
	6.58		*7.31*		*6.52*	
MOM1	-0.511		-0.590		-0.848	*
	-1.22		*-1.77*		*-2.26*	
MOM2	0.536		0.721	*	0.829	*
	1.49		*2.48*		*2.51*	
MOM3	0.701		0.766	*	0.800	*
	1.91		*2.44*		*2.26*	
MOM4	1.014	**	0.955	**	0.608	
	3.04		*3.55*		*1.93*	
Avg R-sqr	0.044		0.029		0.032	
Avg Obs	1577.4		1576.4		1513.3	

Table IX

Results of Monthly Cross-sectional Regressions Using Portfolio Average Price-Impact Parameters: λ_p^K, λ_p^{GH}, λ_p^{FV}, and λ_p^S

This table reports the monthly Fama-MacBeth (1973)-type cross-sectional regressions using the *portfolio average price-impact parameters* for NYSE stocks over the 276 months (23 years: 198301-200512). To obtain the portfolio average price-impact parameters (or lambdas), each month the component stocks are first split into 10 portfolios (with the equal number of stocks) after being sorted in ascending order by firm size (MV) and then each of the 10 portfolios is again split into 10 portfolios after being sorted by each of the four GRT-adjusted price-impact parameters (λ^K, λ^{GH}, λ^{FV}, λ^S), resulting in 100 portfolios for each price-impact parameter in each month. Then the portfolio average lambda is computed for each portfolio and this average value is assigned to the component stocks in the portfolio. λ^K, λ^{GH}, λ^{FV}, and λ^S are defined as in the previous tables and the corresponding portfolio average price-impact parameters are denoted as λ_p^K, λ_p^{GH}, λ_p^{FV}, and λ_p^S, respectively. The dependent variable is FF3EXSRET1 or FF3EXSRET2 in each panel. The definitions of the other variables are as follows: *FF3EXSRET1*: the risk-adjusted excess return (in excess of the risk-free rate proxied by the one-month T-bill rate) using the Fama-French (FF) 3 factors, i.e., the constant term plus the residual from the time-series regression of the excess return on the FF 3 factors using the *entire* sample range of the data; *FF3EXSRET2*: the risk-adjusted excess return using the Fama-French (FF) 3 factors with factor loadings being estimated from the 5-year rolling regressions, i.e., R_i^* computed each month with the current month data from the equation, $R_i^* = (R_i - R_f) - [\hat{\beta}_1 MKT + \hat{\beta}_2 SMB + \hat{\beta}_3 HML]$, after the factor loadings ($\alpha, \beta_1, \beta_2, \beta_3$) are first estimated for *each month* using the time-series data of the past 60 months (at least 24 months) in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \varepsilon$, where R_i, R_f, and R_m are the individual stock return, the risk-free rate, and the market index return, respectively, while *MKT*, *SMB*, and HML are FF 3 factors; *LAM_PAVG*: one of the four portfolio average price-impact parameters (lambdas); *BTM*: natural logarithm of BM_w, which is the Winsorized value (at the 0.5^{th} and 99.5^{th} percentiles) of a book-to-market ratio (BM = BV/MV), where the book value (BV) is defined as common equity plus deferred taxes in $million and the market value (MV) is defined as the previous month-end stock price times the number of shares outstanding (in $million); *MOM1*: compounded holding period return of a stock over the most recent 3 months (from month *t-1* to month *t-3*); *MOM2*: compounded holding period return over the next recent 3 months (from month *t-4* to month *t-6*); *MOM3*: compounded holding period return over the 3 months from month *t-7* to month *t-9*; *MOM4*: compounded holding period return over the 3 months from month *t-9* to month *t-12*. To remove nonstationarity, the Gallant, Rossi, and Tauchen (GRT) (1993) procedure has been applied to λ^K, λ^{GH}, λ^{FV}, and λ^S for each firm over the sample period prior to computing the portfolio averages. BTM is also GRT-adjusted before conducting the cross-sectional regressions (GRT-adjusted variables are indicated by superscript "a"). The values in the first row for each explanatory variable are the time-series averages of coefficients obtained from the month-by-month cross-sectional regressions, and the values italicized in the second row of each variable are *t*-statistics computed based on Fama-MacBeth (1973). The coefficients are all multiplied by 100. *Avg R-sqr* is the average of adjusted R-squared. *Avg Obs* is the monthly average number of companies used in the cross-sectional regressions. The average number of component stocks used each month in the regressions is 1,512.4-1,573.4. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

(Table IX continued)

	Panel A: LAM_PAVG = λ_p^K				Panel B: LAM_PAVG = λ_p^{GH}				Panel C: LAM_PAVG = λ_p^{FV}				Panel D: LAM_PAVG = λ_p^S			
Expla. Variables	FF3EXSRET1		FF3EXSRET2		FF3EXSRET1		FF3EXSRET2		FF3EXSRET1		FF3EXSRET2		FF3EXSRET1		FF3EXSRET2	
Intercept	0.278	**	0.334	**	0.281	**	0.339	**	0.285	**	0.344	**	0.318	**	0.377	**
	2.85		*3.08*		*2.93*		*3.18*		*2.97*		*3.23*		*3.31*		*3.51*	
LAM_PAVG[a]	0.162	**	0.164	**	0.299	**	0.305	**	0.286	**	0.288	**	0.260	**	0.263	**
	4.32		*4.12*		*5.22*		*4.97*		*5.02*		*4.70*		*3.63*		*3.40*	
BTM[a]	0.328	**	0.312	**	0.329	**	0.314	**	0.329	**	0.315	**	0.334	**	0.319	**
	7.48		*6.56*		*7.47*		*6.59*		*7.48*		*6.61*		*7.54*		*6.63*	
MOM1	-0.576		-0.841	*	-0.619		-0.881	*	-0.605		-0.868	*	-0.589		-0.848	*
	-1.72		*-2.24*		*-1.83*		*-2.33*		*-1.79*		*-2.29*		*-1.75*		*-2.25*	
MOM2	0.746	**	0.866	**	0.714	*	0.833	*	0.727	*	0.839	*	0.734	*	0.854	**
	2.58		*2.62*		*2.45*		*2.50*		*2.49*		*2.52*		*2.51*		*2.57*	
MOM3	0.842	**	0.874	*	0.817	**	0.855	*	0.826	**	0.867	*	0.790	*	0.829	*
	2.67		*2.47*		*2.58*		*2.40*		*2.61*		*2.43*		*2.50*		*2.33*	
MOM4	0.933	**	0.579		0.919	**	0.564		0.938	**	0.584		0.944	**	0.594	
	3.48		*1.85*		*3.39*		*1.78*		*3.49*		*1.86*		*3.49*		*1.88*	
Avg R-sqr	0.030		0.033		0.029		0.033		0.029		0.032		0.029		0.032	
Avg Obs	1573.1		1512.9		1573.4		1513.1		1572.7		1512.4		1572.8		1512.5	

Table X
Results of Monthly Cross-sectional Regressions Using Quote Mid-point Returns

This table reports the monthly Fama-MacBeth (1973)-type cross-sectional regressions using quote *mid-point returns* for NYSE stocks. The sample period is: 276 months (23 years: 198301-200512) when the risk-adjustment in returns is performed by the first method; and 252 months (21 years: 198501-200512) when the risk-adjustment in returns is performed by the second method. Quote mid-point returns are calculated using the monthly closing quote mid-point prices after matching the intradaily trades and quotes according to the Lee and Ready (1991) algorithm. The same notations are used for the two dependent variables as before: FF3EXSRET1 and FF3EXSRET2. λ^K, λ^{GH}, λ^{FV}, and λ^S are defined as in the previous tables. The definitions of the other variables are as follows:

FF3EXSRET1: the risk-adjusted excess mid-point return (in excess of the risk-free rate proxied by the one-month T-bill rate) using the Fama-French (FF) 3 factors, i.e., the constant term plus the residual from the time-series regression of the excess mid-point return on the FF 3 factors using the *entire* sample range of the data; *FF3EXSRET2*: the risk-adjusted excess mid-point return using the Fama-French (FF) 3 factors with the factor loadings being estimated from the 5-year rolling regressions, i.e., R_i^* computed each month with the current month data from the equation, $R_i^* = (R_i - R_f) - [\hat{\beta}_1 MKT + \hat{\beta}_2 SMB + \hat{\beta}_3 HML]$, after the factor loadings ($\alpha, \beta_1, \beta_2, \beta_3$) are first estimated for *each month* using the time-series data of the past 60 months (at least 24 months) in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \varepsilon$, where R_i, R_f, and R_m are the individual stock's mid-point return, the risk-free rate, and the market index return, respectively, while *MKT*, *SMB*, and HML are FF 3 factors; *LAM*: one of the four price-impact parameters (lambdas); *BTM*: natural logarithm of BM_w, which is the Winsorized value (at the 0.5th and 99.5th percentiles) of a book-to-market ratio (BM = BV/MV), where the book value (BV) is defined as common equity plus deferred taxes in $million and the market value (MV) is defined as the previous month-end stock price times the number of shares outstanding (in $million); *MOM1*: compounded holding period return of a stock over the most recent 3 months (from month *t-1* to month *t-3*); *MOM2*: compounded holding period return over the next recent 3 months (from month *t-4* to month *t-6*); *MOM3*: compounded holding period return over the 3 months from month *t-7* to month *t-9*; *MOM4*: compounded holding period return over the 3 months from month *t-9* to month *t-12*. To remove nonstationarity, the Gallant, Rossi, and Tauchen (GRT) (1993) procedure has been applied to λ^K, λ^{GH}, λ^{FV}, λ^S, and BTM for each firm over the sample period before conducting the cross-sectional regressions (GRT-adjusted variables are indicated by superscript "a"). The values in the first row for each explanatory variable are the time-series averages of coefficients obtained from the month-by-month cross-sectional regressions, and the values italicized in the second row of each variable are *t*-statistics computed based on Fama-MacBeth (1973). The coefficients are all multiplied by 100. *Avg R-sqr* is the average of adjusted R-squared. *Avg Obs* is the monthly average number of companies used in the cross-sectional regressions. The average number of component stocks used each month in the regressions is 1,467.5-1,558.2. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

(Table X continued)

	Panel A: LAM = λ^{K}		Panel B: LAM = λ^{GH}		Panel C: LAM = λ^{FV}		Panel D: LAM = λ^{S}	
Expla. Variables	FF3EXSRET1	FF3EXSRET2	FF3EXSRET1	FF3EXSRET2	FF3EXSRET1	FF3EXSRET2	FF3EXSRET1	FF3EXSRET2
Intercept	-0.246	-0.397	-0.207	-0.386	-0.227	-0.408	-0.162	-0.370
	-0.66	*-0.88*	*-0.54*	*-0.86*	*-0.60*	*-0.91*	*-0.43*	*-0.84*
LAM[a]	0.267	0.279 **	0.398 *	0.523 **	0.452 **	0.584 **	0.362	0.596 **
	1.83	*2.69*	*2.25*	*3.14*	*2.58*	*3.32*	*1.56*	*2.98*
BTM[a]	0.123	0.130	0.135	0.128	0.134	0.125	0.137	0.126
	0.32	*0.42*	*0.35*	*0.41*	*0.35*	*0.40*	*0.36*	*0.40*
MOM1	0.819	-0.492	0.763	-0.551	0.771	-0.532	0.767	-0.512
	0.56	*-0.38*	*0.52*	*-0.42*	*0.53*	*-0.41*	*0.53*	*-0.39*
MOM2	2.077	-0.003	2.031	-0.072	2.068	-0.044	2.054	-0.013
	0.81	*0.00*	*0.77*	*-0.05*	*0.79*	*-0.03*	*0.79*	*-0.01*
MOM3	-0.072	-0.785	-0.152	-0.829	-0.137	-0.820	-0.164	-0.825
	-0.06	*-0.63*	*-0.13*	*-0.66*	*-0.12*	*-0.66*	*-0.15*	*-0.66*
MOM4	-0.566	1.407	-0.490	1.366	-0.477	1.378	-0.451	1.438
	-0.35	*1.39*	*-0.32*	*1.35*	*-0.31*	*1.37*	*-0.29*	*1.44*
Avg R-sqr	0.008	0.011	0.008	0.011	0.008	0.011	0.008	0.011
Avg Obs	1558.0	1467.7	1558.2	1467.9	1557.7	1467.5	1557.7	1467.5

Table XI
Relations of the Price-Impact Parameters to Alternative Measures

This table reports the monthly average correlations between the four price-impact parameters and other (il)liquidity measures for NYSE stocks over the past 276 months (23 years: 198301-200512. The cross-sectional correlation coefficients are first calculated each month and then the time-series averages of those values over the sample periods are reported here. The definitions of the measures are as follows: λ^K : the price-impact parameter estimated based on Kyle (1985) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5^{th} and 99.5^{th} percentiles; λ^{GH} : the price-impact parameter estimated based on Glosten and Harris (1988) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5^{th} and 99.5^{th} percentiles; λ^{FV} : the price-impact parameter estimated based on Foster and Viswanathan (1993) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5^{th} and 99.5^{th} percentiles; λ^S : the price-impact parameter estimated based on Sadka (2006) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5^{th} and 99.5^{th} percentiles; *Amihud*: the Winsorized (at the 0.5^{th} and 99.5^{th} percentiles) illiquidity measure of Amihud (2002) estimated each month as the average of $|r|/DVOL$, where r is the daily stock return and $DVOL$ is the daily dollar volume in $100,000; *Roll_Gibbs*: the Winsorized (at the 0.5^{th} and 99.5^{th} percentiles) market risk-adjusted effective bid-ask spread of Roll (1984) estimated using the Gibbs sampler, which is of annual frequency obtained from the web site of Joel Hasbrouck; *PS*: the Winsorized (at the 0.5th and 99.5th percentiles) illiquidity measure ($|\gamma|$) of Pastor and Stambaugh (2003), in which γ is estimated from the regression equation, $r^e_{j,d+1,t} = a + br_{j,d,t} + \gamma sign(r^e_{j,d,t})DVOL_{j,d,t} + \varsigma_{j,d+1,t}$, where $r_{j,d,t}$ is the raw return and $r^e_{j,d,t}$ is the excess return (over the CRSP value-weighted index return) of stock j at day d within month t (we require at least 15 days of data per month in the CRSP daily file to estimate γ). The average number of component stocks used in a year is 1,578.1.

	λ^K	λ^{GH}	λ^{FV}	λ^S	Amihud	Roll_Gibbs	PS
λ^K	1						
λ^{GH}	0.936	1					
λ^{FV}	0.931	0.996	1				
λ^S	0.771	0.802	0.803	1			
Amihud	0.539	0.451	0.447	0.374	1		
Roll_Gibbs	0.411	0.286	0.283	0.288	0.626	1	
PS	0.430	0.358	0.355	0.302	0.656	0.477	1

Table XII

A Horse Race with Each of the Three Alternative (Il)liquidity Measures Separately for NYSE Stocks

This table runs a horse race in the monthly Fama-MacBeth (1973)-type cross-sectional regressions for comparison purposes using 3 alternative (il)liquidity measures for NYSE stocks over the past 276 months (23 years: 198301-200512). Each of Panels A-C reports the regression results comparable to Panels A and B in Tables VI-VII using one of the 3 alternative measures: Amihud, Roll_Gibbs, and TURN. The dependent variable is FF3EXSRET1 or FF3EXSRET2 in each panel. The definitions of the variables are as follows: *FF3EXSRET1*: the risk-adjusted excess return (in excess of the risk-free rate proxied by the one-month T-bill rate) using the Fama-French (FF) 3 factors, i.e., the constant term plus the residual from the time-series regression of the excess return on the FF 3 factors using the *entire* sample range of the data; *FF3EXSRET2*: the risk-adjusted excess return using the Fama-French (FF) 3 factors with factor loadings being estimated from the 5-year rolling regressions, i.e., R_i^* computed each month with the current month data from the equation, $R_i^* = (R_i - R_f) - [\hat{\beta}_1 MKT + \hat{\beta}_2 SMB + \hat{\beta}_3 HML]$, after the factor loadings ($\alpha, \beta_1, \beta_2, \beta_3$) are first estimated for *each month* using the time-series data of the past 60 months (at least 24 months) in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \varepsilon$, where R_i, R_f, and R_m are the individual stock return, the risk-free rate, and the market index return, respectively, while *MKT*, *SMB*, and HML are FF 3 factors; *ALT*: one of the three alternative (il)liquidity measures; *Amihud*: the Winsorized (at the 0.5th and 99.5th percentiles) illiquidity measure of Amihud (2002) estimated each month as the average of |*r*|/*DVOL*, where *r* is the daily stock return and *DVOL* is the daily dollar volume in $100,000; *Roll_Gibbs*: the Winsorized (at the 0.5th and 99.5th percentiles) market risk-adjusted effective bid-ask spread of Roll (1984) estimated using the Gibbs sampler, which is of annual frequency obtained from the web site of Joel Hasbrouck; *PS*: the Winsorized (at the 0.5th and 99.5th percentiles) illiquidity measure ($|\gamma|$) of Pastor and Stambaugh (2003), in which γ is estimated from the regression equation, $r^e_{j,d+1,t} = a + br_{j,d,t} + \gamma sign(r^e_{j,d,t}) DVOL_{j,d,t} + \varsigma_{j,d+1,t}$, where $r_{j,d,t}$ is the raw return and $r^e_{j,d,t}$ is the excess return (over the CRSP value-weighted index return) of stock *j* at day *d* within month *t* (we require at least 15 days of data per month in the CRSP daily file to estimate γ); *BTM*: natural logarithm of BM_w, which is the Winsorized value (at the 0.5^{th} and 99.5^{th} percentiles) of a book-to-market ratio (BM = BV/MV), where the book value (BV) is defined as common equity plus deferred taxes in $million and the market value (MV) is defined as the previous month-end stock price times the number of shares outstanding (in $million); *MOM1*: compounded holding period return of a stock over the most recent 3 months (from month *t-1* to month *t-3*); *MOM2*: compounded holding period return over the next recent 3 months (from month *t-4* to month *t-6*); *MOM3*: compounded holding period return over the 3 months from month *t-7* to month *t-9*; *MOM4*: compounded holding period return over the 3 months from month *t-9* to month *t-12*. For monthly regressions, we keep the annual values of Roll_Gibbs constant over the 12 months within each year. To remove nonstationarity, the Gallant, Rossi, and Tauchen (GRT) (1993) procedure has been applied to Amihud, Roll_Gibbs, PS, and BTM for each firm over the sample period before conducting the cross-sectional regressions (GRT-adjusted variables are indicated by superscript “a”). The values in the first row for each explanatory variable are the time-series averages of coefficients obtained from the month-by-month cross-sectional regressions, and the values italicized in the second row of each variable are t-statistics computed based on Fama-MacBeth (1973). The coefficients are all multiplied by 100. *Avg R-sqr* is the average of adjusted R-squared. *Avg Obs* is the monthly average number of companies used in the cross-sectional regressions. The average number of component stocks used each month in the regressions is 1,197.4-1,577.0. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

(Table XII continued)

Expla. Variables	Panel A: ALT = Amihud FF3EXSRET1		Panel A: ALT = Amihud FF3EXSRET2		Panel B: ALT = Roll_Gibbs FF3EXSRET1		Panel B: ALT = Roll_Gibbs FF3EXSRET2		Panel C: ALT = PS FF3EXSRET1		Panel C: ALT = PS FF3EXSRET2	
Intercept	0.339	**	0.394	**	0.404	**	0.473	**	0.383	**	0.443	**
	3.62		*3.76*		*3.38*		*3.81*		*4.05*		*4.19*	
ALT[a]	4.723	**	5.352	**	-1.458		-3.903		1.014		1.046	
	2.98		*3.50*		*-0.12*		*-0.31*		*1.69*		*1.67*	
BTM[a]	0.323	**	0.309	**	0.417	**	0.393	**	0.334	**	0.322	**
	7.39		*6.54*		*8.67*		*7.59*		*7.59*		*6.78*	
MOM1	-0.462		-0.727		-1.164	**	-1.476	**	-0.533		-0.809	*
	-1.38		*-1.94*		*-3.32*		*-3.73*		*-1.58*		*-2.14*	
MOM2	0.809	**	0.926	**	0.421		0.491		0.703	*	0.811	*
	2.79		*2.80*		*1.40*		*1.44*		*2.40*		*2.45*	
MOM3	0.875	**	0.911	*	0.362		0.435		0.756	*	0.799	*
	2.76		*2.56*		*1.17*		*1.30*		*2.38*		*2.23*	
MOM4	1.001	**	0.663	*	0.736	**	0.424		0.909	**	0.558	
	3.72		*2.11*		*2.83*		*1.36*		*3.38*		*1.77*	
Avg R-sqr	0.031		0.034		0.031		0.034		0.029		0.033	
Avg Obs	1576.8		1513.8		1234.7		1197.4		1577.0		1514.0	

Table XIII
A Horse Race with All the (Il)liquidity Measures Together for NYSE Stocks

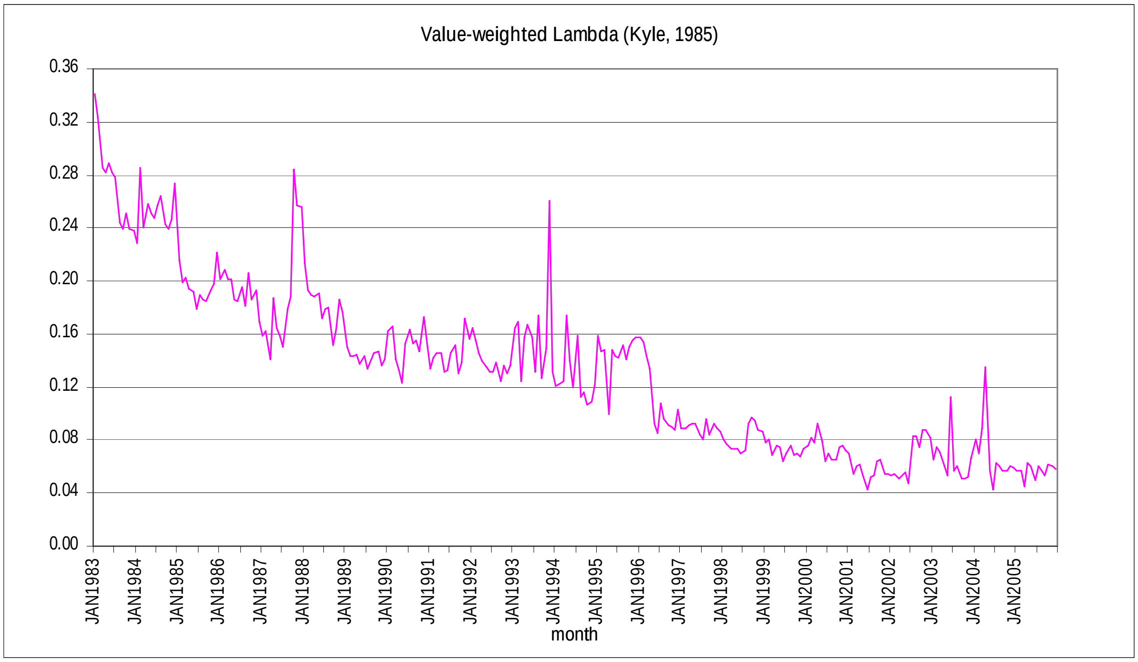
This table runs a horse race in the monthly Fama-MacBeth (1973)-type cross-sectional regressions using one of the four price-impact parameters (λ^K , λ^{GH} , λ^{FV} , and λ^S) together with the 3 alternative (il)liquidity measures for NYSE stocks over the past 276 months (23 years: 198301-200512). λ^K , λ^{GH} , λ^{FV} , and λ^S are defined as in the previous tables. The dependent variable is FF3EXSRET1 or FF3EXSRET2 in each panel. The definitions of the other variables are as follows: *FF3EXSRET1*: the risk-adjusted excess return (in excess of the risk-free rate proxied by the one-month T-bill rate) using the Fama-French (FF) 3 factors, i.e., the constant term plus the residual from the time-series regression of the excess return on the FF 3 factors using the *entire* sample range of the data; *FF3EXSRET2*: the risk-adjusted excess return using the Fama-French (FF) 3 factors with factor loadings being estimated from the 5-year rolling regressions, i.e., R_i^* computed each month with the current month data from the equation, $R_i^* = (R_i - R_f) - [\hat{\beta}_1 MKT + \hat{\beta}_2 SMB + \hat{\beta}_3 HML]$, after the factor loadings $(\alpha, \beta_1, \beta_2, \beta_3)$ are first estimated for *each month* using the time-series data of the past 60 months (at least 24 months) in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \varepsilon$, where R_i, R_f , and R_m are the individual stock return, the risk-free rate, and the market index return, respectively, while *MKT*, *SMB*, and HML are FF 3 factors; *LAM*: one of the four price-impact parameters (lambdas); *Amihud*: the Winsorized (at the 0.5th and 99.5th percentiles) illiquidity measure of Amihud (2002) estimated each month as the average of |*r*|/*DVOL*, where *r* is the daily stock return and *DVOL* is the daily dollar volume in $100,000; *Roll_Gibbs*: the Winsorized (at the 0.5th and 99.5th percentiles) market risk-adjusted effective bid-ask spread of Roll (1984) estimated using the Gibbs sampler, which is of annual frequency obtained from the web site of Joel Hasbrouck; *PS*: the Winsorized (at the 0.5th and 99.5th percentiles) illiquidity measure ($|\gamma|$) of Pastor and Stambaugh (2003), in which γ is estimated from the regression equation, $r^e_{j,d+1,t} = a + br_{j,d,t} + \gamma sign(r^e_{j,d,t}) DVOL_{j,d,t} + \varsigma_{j,d+1,t}$, where $r_{j,d,t}$ is the raw return and $r^e_{j,d,t}$ is the excess return (over the CRSP value-weighted index return) of stock *j* at day *d* within month *t* (we require at least 15 days of data per month in the CRSP daily file to estimate γ); *BTM*: natural logarithm of BM_w, which is the Winsorized value (at the 0.5^{th} and 99.5^{th} percentiles) of a book-to-market ratio (BM = BV/MV), where the book value (BV) is defined as common equity plus deferred taxes in $million and the market value (MV) is defined as the previous month-end stock price times the number of shares outstanding (in $million); *MOM1*: compounded holding period return of a stock over the most recent 3 months (from month *t-1* to month *t-3*); *MOM2*: compounded holding period return over the next recent 3 months (from month *t-4* to month *t-6*); *MOM3*: compounded holding period return over the 3 months from month *t-7* to month *t-9*; *MOM4*: compounded holding period return over the 3 months from month *t-9* to month *t-12*. For monthly regressions, we keep the annual values of Roll_Gibbs constant over the 12 months within each year. To remove nonstationarity, the Gallant, Rossi, and Tauchen (GRT) (1993) procedure has been applied to λ^K , λ^{GH} , λ^{FV} , λ^S , Amihud, Roll_Gibbs, PS, and BTM for each firm over the sample period before conducting the cross-sectional regressions (GRT-adjusted variables are indicated by superscript "a"). The values in the first row for each explanatory variable are the time-series averages of coefficients obtained from the month-by-month cross-sectional regressions, and the values italicized in the second row of each variable are t-statistics computed based on Fama-MacBeth (1973). The coefficients are all multiplied by 100. *Avg R-sqr* is the average of adjusted R-squared. *Avg Obs* is the monthly average number of companies used in the cross-sectional regressions. The average number of component stocks used each month in the regressions is 1,198.9-1,236.2. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

(Table XIII continued)

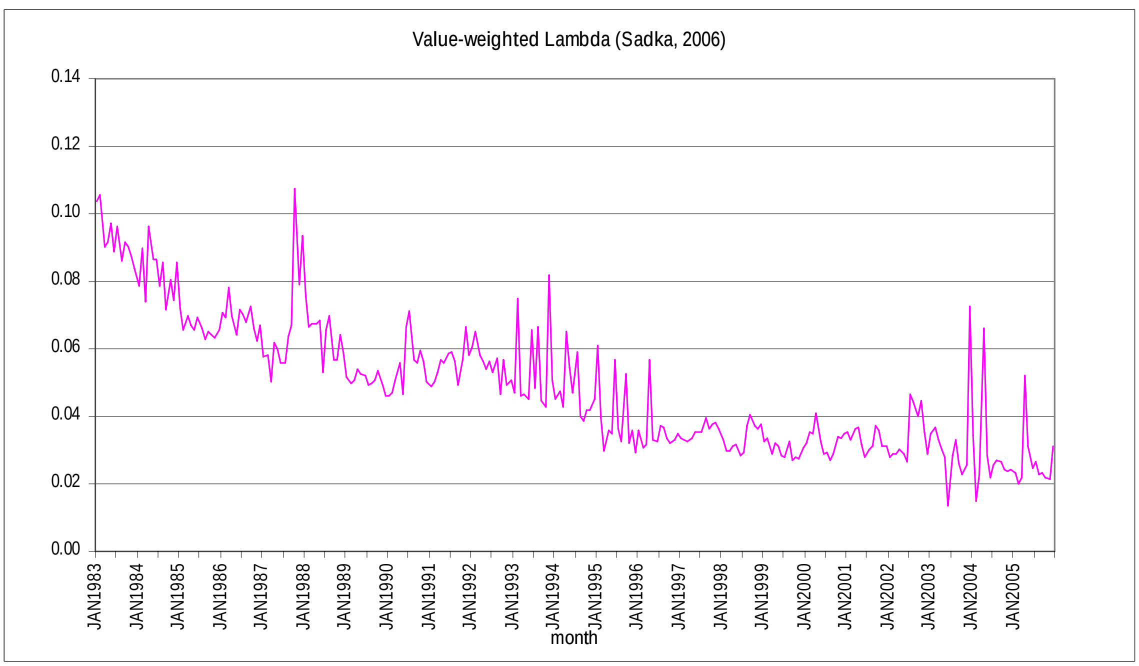
	Panel A: LAM = λ^{K}		Panel B: LAM = λ^{GH}		Panel C: LAM = λ^{FV}		Panel D: LAM = λ^{S}	
Expla. Variables	FF3EXSRET1	FF3EXSRET2	FF3EXSRET1	FF3EXSRET2	FF3EXSRET1	FF3EXSRET2	FF3EXSRET1	FF3EXSRET2
Intercept	0.364 **	0.431 **	0.347 **	0.421 **	0.352 **	0.424 **	0.374 **	0.443 **
	2.97	*3.38*	*2.86*	*3.32*	*2.89*	*3.34*	*3.06*	*3.45*
LAM[a]	0.097 **	0.094 *	0.210 **	0.196 **	0.207 **	0.193 **	0.208 **	0.201 **
	2.70	*2.55*	*3.85*	*3.57*	*3.77*	*3.50*	*3.20*	*2.93*
Amihud[a]	*4.307* **	*4.995* **	*3.750* *	*4.644* **	*3.702* *	*4.488* **	*4.458* **	*5.203* **
	3.14	*3.29*	*2.36*	*2.71*	*2.29*	*2.60*	*3.13*	*3.37*
Roll_Gibbs[a]	*-20.429*	*-22.664*	*-17.723*	*-20.689*	*-17.866*	*-20.267*	*-19.501*	*-22.142*
	-1.67	*-1.73*	*-1.42*	*-1.56*	*-1.43*	*-1.53*	*-1.58*	*-1.69*
PS[a]	*0.729*	*0.770*	*0.911*	*0.883*	*0.916*	*0.890*	*0.779*	*0.762*
	0.60	*0.61*	*0.66*	*0.64*	*0.66*	*0.64*	*0.72*	*0.69*
BTM[a]	0.385 **	0.370 **	0.384 **	0.370 **	0.389 **	0.375 **	0.387 **	0.373 **
	7.77	*7.00*	*7.77*	*7.04*	*7.88*	*7.13*	*7.82*	*7.05*
MOM1	-1.153 **	-1.460 **	-1.150 **	-1.453 **	-1.160 **	-1.467 **	-1.127 **	-1.438 **
	-3.31	*-3.73*	*-3.31*	*-3.71*	*-3.33*	*-3.75*	*-3.24*	*-3.68*
MOM2	0.493	0.550	0.476	0.534	0.466	0.518	0.491	0.546
	1.64	*1.61*	*1.59*	*1.57*	*1.55*	*1.52*	*1.64*	*1.60*
MOM3	0.286	0.301	0.292	0.305	0.297	0.310	0.285	0.301
	0.93	*0.87*	*0.95*	*0.87*	*0.96*	*0.89*	*0.93*	*0.87*
MOM4	0.708 **	0.403	0.691 **	0.395	0.687 **	0.393	0.727 **	0.430
	2.72	*1.29*	*2.65*	*1.27*	*2.63*	*1.26*	*2.78*	*1.37*
Avg R-sqr	0.042	0.046	0.042	0.046	0.042	0.046	0.042	0.046
Avg Obs	1236.2	1199.0	1236.2	1199.0	1236.1	1198.9	1236.2	1199.0

Figure 1. Trends of the Value-Weighted Price-Impact Parameters for NYSE Stocks

The following graphs show the trends of the price-impact parameters for NYSE stocks over the past 276 months (23 years: 198301-200512). Each of the two graphs is a time-series plot of the monthly cross-sectional [market-value (MV) weighted] average of the price-impact parameters over the sample period. The price-impact parameters of individual stocks are all multiplied by 10^6 and then Winsorized each month at the 0.5th and 99.5th percentiles before computing the MV-weighted cross-sectional averages. Figure 1(a) is for the price-impact parameter estimated based on Kyle (1985), λ^K and Figure 1(b) is for that based on Sadka (2006), λ^S. The average number of component stocks used in each month is 1,587.1.



(a)



(b)

Table A1

Results of Monthly Cross-sectional Regressions: Different Specifications With or Without SIZE

This table reports the four different specifications of the monthly Fama-MacBeth (1973)-type cross-sectional regressions with or without SIZE for NYSE stocks over the 276 months (23 years: 198301-200512). In Panels A, B, C, and D, the price-impact parameter used is the one estimated based on Kyle (1985, λ^{K}), Glosten and Harris (1988, λ^{GH}), Foster and Viswanathan (1993, λ^{FV}), and Sadka (2006, λ^{S}), respectively. The dependent variable is FF3EXSRET1 or FF3EXSRET2 in each panel. The definitions of the variables are as follows: *FF3EXSRET1*: the risk-adjusted excess return using the Fama-French (FF) 3 factors, i.e., the constant term plus the residual from the time-series regression of the excess return on the FF 3 factors using the *entire* sample range of the data; *FF3EXSRET2*: the risk-adjusted excess return using the Fama-French (FF) 3 factors with factor loadings being estimated from the 5-year rolling regressions, i.e., R_i^* computed each month with the current month data from the equation, $R_i^* = (R_i - R_f) - [\hat{\beta}_1 MKT + \hat{\beta}_2 SMB + \hat{\beta}_3 HML]$, after the factor loadings ($\alpha, \beta_1, \beta_2, \beta_3$) are first estimated for *each month* using the time-series data of the past 60 months (at least 24 months) in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \varepsilon$, where R_i, R_f, and R_m are the individual stock return, the risk-free rate, and the market index return, respectively, while *MKT*, *SMB*, and HML are FF 3 factors; λ^{K} : the price-impact parameter estimated based on Kyle (1985) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5th and 99.5th percentiles; λ^{GH} : the price-impact parameter estimated based on Glosten and Harris (1988) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5th and 99.5th percentiles; λ^{FV} : the price-impact parameter estimated based on Foster and Viswanathan (1993) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5th and 99.5th percentiles; λ^{S} : the price-impact parameter estimated based on Sadka (2006) using intradaily dollar order flows available within each month, multiplied by 10^6, and then Winsorized at the 0.5th and 99.5th percentiles; *SIZE*: natural logarithm of MV, where MV is the market value defined as the previous month-end stock price times the number of shares outstanding (in $million); *BTM*: natural logarithm of BM_w, which is the Winsorized value (at the 0.5th and 99.5th percentiles) of a book-to-market ratio (BM = BV/MV), where the book value (BV) is defined as common equity plus deferred taxes in $million and the market value (MV) is defined as the previous month-end stock price times the number of shares outstanding (in $million); *MOM1*: compounded holding period return of a stock over the most recent 3 months (from month *t-1* to month *t-3*); *MOM2*: compounded holding period return over the next recent 3 months (from month *t-4* to month *t-6*); *MOM3*: compounded holding period return over the 3 months from month *t-7* to month *t-9*; *MOM4*: compounded holding period return over the 3 months from month *t-9* to month *t-12*. To remove nonstationarity, the Gallant, Rossi, and Tauchen (GRT) (1993) procedure has been applied to λ^{K}, λ^{GH}, λ^{FV}, λ^{S}, SIZE, and BTM for each firm over the sample period before conducting the cross-sectional regressions (GRT-adjusted variables are indicated by superscript "a"). The values in the first row for each explanatory variable are the time-series averages of coefficients obtained from the month-by-month cross-sectional regressions, and the values italicized in the second row of each variable are *t*-statistics computed based on Fama-MacBeth (1973). The coefficients are all multiplied by 100. *Avg R-sqr* is the average of adjusted R-squared. *Avg Obs* is the monthly average number of companies used in the cross-sectional regressions. The average number of component stocks used each month in the regressions is 1,505.8-1,598.3. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

(Table A1 continued: Panels A and B)

Panel A: with Lambdas Based on Kyle (1985)

	Dep Var = FF3EXSRET1							
Expla. Variables	1		2		3		4	
Intercept	0.232	**	1.447	**	2.668	**	2.860	**
	2.99		*6.60*		*12.23*		*14.09*	
$\lambda^{K,a}$	0.167	**			-0.395	**	-0.386	**
	4.34				*-12.06*		*-11.48*	
SIZE[a]			-0.200	**	-0.344	**	-0.383	**
			-6.84		*-12.45*		*-16.41*	
BTM[a]							0.094	*
							2.10	
MOM1							-0.178	
							-0.52	
MOM2							1.253	**
							4.18	
MOM3							1.402	**
							4.47	
MOM4							1.423	**
							5.16	
Avg R-sqr	0.005		0.005		0.009		0.036	
Avg Obs	1597.7		1521.2		1520.7		1506.0	

Panel B: with Lambdas Based on Glosten and Harris (1988)

	Dep Var = FF3EXSRET1							
Expla. Variables	1		2		3		4	
Intercept	0.225	**	1.447	**	2.189	**	2.431	**
	2.96		*6.60*		*9.86*		*12.16*	
$\lambda^{GH,a}$	0.333	**			-0.484	**	-0.492	**
	6.04				*-10.72*		*-10.69*	
SIZE[a]			-0.200	**	-0.287	**	-0.333	**
			-6.84		*-9.97*		*-14.22*	
BTM[a]							0.102	*
							2.30	
MOM1							-0.157	
							-0.46	
MOM2							1.296	**
							4.32	
MOM3							1.465	**
							4.70	
MOM4							1.465	**
							5.32	
Avg R-sqr	0.004		0.005		0.007		0.035	
Avg Obs	1598.3		1521.2		1521.2		1506.1	

(Table A1 continued: Panels C and D)

Panel C: with Lambdas Based on Foster and Viswanathan (1993)

	Dep Var = FF3EXSRET1							
Expla. Variables	1		2		3		4	
Intercept	0.229	**	1.447	**	2.215	**	2.450	**
	3.01		*6.60*		*9.89*		*12.23*	
$\lambda^{FV,a}$	0.313	**			-0.501	**	-0.507	**
	5.59				*-11.27*		*-11.18*	
SIZE[a]			-0.200	**	-0.290	**	-0.335	**
			-6.84		*-10.01*		*-14.28*	
BTM[a]							0.103	*
							2.31	
MOM1							-0.163	
							-0.47	
MOM2							1.302	**
							4.34	
MOM3							1.466	**
							4.70	
MOM4							1.469	**
							5.35	
Avg R-sqr	0.004		0.005		0.007		0.035	
Avg Obs	1597.1		1521.2		1520.7		1505.8	

Panel D: with Lambdas Based on Sadka (2006)

	Dep Var = FF3EXSRET1							
Expla. Variables	1		2		3		4	
Intercept	0.271	**	1.447	**	2.188	**	2.384	**
	3.59		*6.60*		*10.17*		*12.43*	
$\lambda^{S,a}$	0.276	**			-0.706	**	*-0.68*	**
	3.87				*-6.37*		*-6.55*	
SIZE[a]			-0.200	**	-0.286	**	-0.326	**
			-6.84		*-10.23*		*-14.42*	
BTM[a]							0.109	*
							2.50	
MOM1							-0.193	
							-0.56	
MOM2							1.246	**
							4.17	
MOM3							1.426	**
							4.60	
MOM4							1.437	**
							5.18	
Avg R-sqr	0.003		0.005		0.007		0.034	
Avg Obs	1597.2		1521.2		1520.6		1505.8	