Theory-Based Illiquidity and Asset Pricing

Tarun Chordia
Goizueta Business School, Emory University

Sahn-Wook Huh
School of Management, State University of New York (SUNY) at Buffalo

Avanidhar Subrahmanyam
Anderson School, University of California

Many proxies of illiquidity have been used in the literature that relates illiquidity to asset prices. These proxies have been motivated from an empirical standpoint. In this study, we approach liquidity estimation from a theoretical perspective. Our method explicitly recognizes the analytic dependence of illiquidity on more primitive drivers such as trading activity and information asymmetry. More specifically, we estimate illiquidity using structural formulae in line with Kyle's (1985) lambda for a comprehensive sample of stocks. The empirical results provide evidence that theory-based estimates of illiquidity are priced in the cross-section of expected stock returns, even after accounting for risk factors, firm characteristics known to influence returns, and other illiquidity proxies prevalent in the literature. (*JEL* G12, G14)

The question of whether investors demand higher returns from less liquid securities is an enduring one in financial economics. In a seminal paper, Amihud and Mendelson (1986) find evidence that asset 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

We gratefully appreciate the constructive and thoughtful comments of an anonymous referee and Matthew Spiegel (the editor). We also thank Antonio Bernardo, Michael Brennan, Susan Christoffersen, Kee H. Chung, Na Dai, Vihang Errunza, Ruslan Goyenko, Hemantha Herath, Paul Irvine, Adam Kolasinski, Jiro Kondo, Thomas McInish, Emmanuel Morales-Camargo, Amrita Nain, Joseph Ogden, Yun Park, Christo Pirinsky, Unyong Pyo, John Schatzberg, Lawrence Southwick Jr., Jaeyoung Sung, Tony Tang, Cristian Tiu, Samir Trabelsi, Gautam Vora, Bob Welch, Dong-Chul Won, and seminar participants at Yale University, SUNY at Buffalo, Ajou University, Brock University, McGill University, California State University at Fullerton, University of New Mexico, the 2006 Annual Conference on Market Structure and Market Integrity, the 2007 KAFA–KFAs Joint Conference, the 2007 Financial Management Association Meetings, and the 2008 CICF Conference for valuable feedback. Research assistance was ably provided by Mi-Ae Kim. Huh gratefully acknowledges 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 authors' responsibility. Send correspondence to A. Subrahmanyam, Anderson School of Management, University of California at Los Angeles, Los Angeles, CA; telephone: (310)-825-5355; fax: (310)-206-5455; Cell: 90095-1481. E-mail: subra@anderson.ucla.edu.

doi:10.1093/rfs/hhn121 Advance Access publication January 27, 2009

role of liquidity as a determinant of expected returns. Furthermore, Pástor and Stambaugh (2003) and Acharya and Pedersen (2005) relate liquidity risk to expected stock returns.[1]

An important issue in studies that relate illiquidity to asset prices is the measurement of illiquidity. Other than direct empirical measurements of illiquidity by the bid-ask spread, the approach taken in the literature has 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), based on the analysis of Glosten and Harris (1988), 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 estimates of these and other measures including the Roll (1984) measure.

While these empirical proxies have added considerably to our understanding of illiquidity, there are some concerns. First, because these measures have yielded mixed results, the significance of the findings is difficult to interpret. For instance, at odds with the liquidity premium argument, Brennan and Subrahmanyam (1996) find a negative relation between bid-ask spread and expected returns, and Spiegel and Wang (2005) do not find a significant relation between expected returns and empirically motivated measures of illiquidity. Second, the empirical arguments proposed to justify liquidity proxies are not motivated by theory. Although many microstructure theories have been developed, extant economic models are unable to map precisely onto the Amihud (2002) construct of the ratio of absolute return to volume. Third, illiquidity is endogenous and depends on many variables that could be related to asset prices. For instance, illiquidity depends on volatility, but volatility may be related to expected returns via traditional risk-return arguments.

In this paper, we propose a new approach for measuring illiquidity and relate our measures to expected asset returns, thereby providing stronger theoretical underpinnings to the empirical illiquidity-return relation than the existing literature. Specifically, we turn to theory in order to consider illiquidity estimates that can be estimated by way of closed-form expressions. The basis for our work emanates from Brennan and Subrahmanyam (1995), who test a structural representation of a theoretically derived estimate of illiquidity and relate it to analyst following. Their estimates derive from the price-impact measure, lambda, which, in turn, is based on the Kyle (1985) model and its adaptation by Admati and Pfleiderer (1988) to explain intraday patterns. The advantage

[1] Two recent theoretical papers attempt to endogenize liquidity in asset-pricing settings. Eisfeldt (2004) relates liquidity to 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. He shows that such a measure may dynamically vary with market returns and, hence, help provide a rationale for liquidity dynamics documented in the literature.

of estimating the equilibrium versions of Kyle lambdas is that the expressions are in terms of quantities that are relatively easy to comprehend and for which plausible empirical proxies can be devised at low costs.

Our analysis considers illiquidity to be endogenous insofar as it arises as an outcome of trading patterns in financial markets. Unlike a stock's return beta, which depends on influences extraneous to financial markets, such as a firm's line of business and the cyclicality of a firm's revenue stream, the endogeneity of illiquidity makes it difficult to interpret results from asset-pricing regressions. For example, illiquidity depends on total volatility, including systematic risk.[2] Lack of adequate controls for systematic risk (e.g., only through the CAPM, as in Amihud and Mendelson, 1986; Brennan and Wang, 2006) could create the appearance that illiquidity is priced because illiquidity depends on true systematic risk. Liquidity also depends on volume. If volume captures investor sentiment (viz., Baker and Stein, 2004), then illiquidity may again appear to be priced, even though what the researcher may be capturing is the impact of volume through its impact on illiquidity. As such, a complete approach to understanding the pricing of illiquidity would model illiquidity's dependence on primitive economic forces and separately control for systematic risk and trading volume, which is what our study attempts to accomplish. Furthermore, in contrast to the *ad hoc* measures of illiquidity in the literature, the functional form of illiquidity that we use is obtained from an equilibrium setting.

We estimate two variants of closed-form expressions for Kyle lambdas, one of which assumes perfect information signals, while the other postulates noisy signals. Many of the empirical proxies for the inputs to the Kyle (1985) model are similar to those used by Brennan and Subrahmanyam (1995). In examining the time-series behavior of such lambdas, we find a decline in the measures over time, which mirrors the behavior of other illiquidity proxies, such as bid-ask spreads (Jones, 2002). We then examine whether these lambdas are priced in the cross-section of stock returns, using a comprehensive set of NYSE/AMEX (and NASDAQ) stocks over the last twenty-seven years. After controlling for known characteristics such as firm size, book-to-market equity, and momentum as well as for known sources of risk such as the Fama and French (1993) factors, we find evidence that both versions of Kyle lambdas are priced in the cross-section. The evidence of illiquidity pricing is stronger for smaller firms. We check the robustness of our results using quote midpoint returns. As a further robustness check, we run a "horse race" with other commonly used (il)liquidity measures, demonstrating that theory-based illiquidity is priced even after accounting for the effects of other competing (il)liquidity measures.

The remainder of this paper is organized as follows. In Section 1, we present the theoretical background and estimation of the two theory-based illiquidity measures in the context of Kyle lambdas. Section 2 describes the methodology.

[2] In strategic trading models, this holds as long as agents have private information about firm-specific as well as systematic components of firm value, as in Brennan, Jegadeesh, and Swaminathan (1993); Subrahmanyam (1991); or Kumar and Seppi (1993). For empirical evidence, see Chordia, Roll, and Subrahmanyam (2000).

Section 3 outlines the data, definitions, descriptive statistics, and adjustments for nonstationarity in variables. Section 4 discusses the empirical results and robustness checks. In Section 5, we compare the effects of the theory-based measures with those of other alternative (il)liquidity measures and show that illiquidity pricing is not driven mainly by a few outlier stocks. Section 6 concludes.

1. Estimates of Kyle Lambdas

In this section, we provide the theoretical background for our lambda estimation. We begin by linking illiquidity and asset pricing in the context of Kyle lambdas.

1.1 The link between Kyle lambdas and illiquidity pricing

The Kyle (1985) model does not provide a direct link to illiquidity pricing. However, assuming that a liquidity trader is the marginal agent allows the incorporation of a link. We consider the one-shot version of the Kyle (1985) setting and suppose that an asset is traded over two dates. At date 2, it pays off $\widetilde{W} = \overline{W} + \delta$, where $\overline{W}$ is nonstochastic and the payoff innovation (δ) is a normally distributed variable with a mean of zero. Informed traders obtain a (possibly noisy) signal about δ. Trading of the asset occurs at date 1. As usual, the price P is set to be of the form $P = \overline{W} + \lambda Q$, where λ is the impact of order flows on prices, and Q is the total order flow. Prior to the date 1 trading (at date 0), a "discretionary" uninformed (liquidity) trader contemplates investing in the asset. This trader's demand is denoted by D, which is normally distributed with its mean zero. In addition, there is a set of nondiscretionary liquidity traders whose total demand, denoted by U, is also normally distributed with a zero mean. D and U are independent of each other. We define $z \equiv U + D$ to be the total demand from the liquidity traders. The discretionary and nondiscretionary liquidity demands are independent of δ. For a given λ, the expected cost to the discretionary liquidity trader is then given by $E[(P - \widetilde{W})D] = \lambda D^2$.

Assuming that the risk-free discount rate is zero across dates 0 and 2, we normalize the asset's supply to one share.[3] The date 0 price is the shadow price at which the discretionary trader is indifferent between holding the stock and not doing so. At date 0, the risk-neutral discretionary liquidity trader will be willing to pay an amount

$$\overline{W} - \lambda D^2.$$

Thus, the expected price change across dates 0 and 2 is given by λD^2, and is thus proportional to λ (ignoring cross-sectional variation in D).[4] It follows that

[3] The supply of shares does not play any role at date 1 because prices are set by risk-neutral market makers who are willing to absorb any quantity of excess shares at an unbiased price.

[4] Note that the expected price change across dates 1 and 2 is zero (the date 1 price is semistrong efficient), and the expected price change across dates 0 and 1 is equal to the expected price change across dates 0 and 2. The

expected future returns are linearly related to λ divided by the initial price of the stock. As we describe below, for our empirical work we estimate λ's each month for each stock and proxy the initial price by the market price of the stock (P) as of the end of the month previous to the one in which λ is measured.

Our study uses structural estimates of two versions of Kyle lambdas: one with signal noise and the other without it. In the following two subsections, we present the theory-based illiquidity measures, and discuss how to estimate them using proxy variables. We first present our main results using a base set of inputs, and then examine in Subsection 4.3 the robustness of our results to using: (i) midpoint returns (i.e., returns estimated using the midpoint of bid and ask quotes); and (ii) alternative input variables.

1.2 An illiquidity measure without noise in the signals

When the informed traders observe, without any noise, a signal that is informative about the payoff on a risky asset, the Appendix shows in detail that the illiquidity (or price impact) measure, lambda, in a standard Kyle (1985) market is given by

$$\lambda = \frac{\sqrt{N v_\delta}}{(N+1)\sqrt{v_z}}, \tag{1}$$

where N is the number of informed traders, v_δ is the variance of the payoff, and v_z is the variance of uninformed trades. Dividing both sides of Equation (1) by price P in order to get a price-scaled illiquidity measure, we have[5]

$$\frac{\lambda}{P} = P^{-1}\frac{\sqrt{N v_\delta}}{(N+1)\sqrt{v_z}} = \frac{\sqrt{N}\sqrt{Var(R)}}{(N+1)\sqrt{v_z}} = \frac{N^{0.5} std(R)}{(N+1) std(z)}, \tag{2}$$

where R is the asset return and $std(z)$ is the standard deviation of uninformed trades. Equation (2) is our first measure of illiquidity used in this study, and we call it *ILLIQ_1*.

To estimate *ILLIQ_1* each month for each stock, we employ proxy variables as inputs for each of the original variables in Equation (2). Our approach in this subsection is not to condition on any specific source of information (such as earnings), but to assume that private information is about value innovations, as reflected in the series of stock price movements. We use analyst following to proxy for informed agents.[6] This approach toward estimating lambdas is very

expected price change across dates 0 and 2 is, therefore, the only unique, nonzero expected price change in our model.

[5] For Equation (2), note that at time t the conditional variance of returns is $Var(R) = Var(\frac{P_t}{P_{t-1}} - 1 \mid I_{t-1}) = \frac{Var(P_t)}{P_{t-1}^2}$.

[6] Information on analyst opinions, of course, eventually becomes publicly available. Green (2006) shows, however, that analysts often provide information privately to preferred clients, and that analysts' revisions have significant profit potential, which is consistent with these agents producing private information.

similar to that of Brennan and Subrahmanyam (1995). Our specific inputs are as follows:

N: One plus the number of analysts following a firm in each month, notated as *ANA*.[7]

$std(R)$: The standard deviation of daily returns within the previous month (month $t-1$), notated as *STD*(*RET*). To obtain this variable for each month, we use firms that have at least ten daily returns in the previous month from the CRSP daily file.

$std(z)$: The average of daily dollar volume (in million dollars) within the previous month, notated as AVG(DVOL).[8] To obtain this variable for each month, we use firms that have at least ten daily trading records in the previous month from the CRSP daily file.

We now discuss how to measure lambdas when information signals are assumed to be noisy and diverse.

1.3 An illiquidity measure with noisy information signals

When the informed traders observe diverse signals so that each trader observes the asset's payoff plus an error term that is independent and identically distributed across agents, the Appendix shows that Kyle's (1985) measure, λ, is given by

$$\lambda = \frac{v_\delta}{(N+1)v_\delta + 2v_\varepsilon}\sqrt{\frac{N(v_\delta + v_\varepsilon)}{v_z}}, \tag{3}$$

where N is the number of informed traders, v_δ is the variance of the asset payoff, v_z is the variance of uninformed trades, and v_ε is the variance of signal innovations. Dividing both sides of Equation (3) by P_{t-1}, we have

$$\frac{\lambda}{P} = P^{-1}\frac{v_\delta}{(N+1)v_\delta + 2v_\varepsilon}\sqrt{\frac{N(v_\delta + v_\varepsilon)}{v_z}}. \tag{4}$$

Equation (4) is our second measure of illiquidity used in this study, and we call it *ILLIQ_2*. Note that this measure requires a proxy for the variance of the signal noise, as well as that of the signal itself. It is difficult to obtain such a proxy from the return series alone. Therefore, in an approach different from that in the previous subsection, we condition on a specific informational event, namely, earnings announcements, in computing the lambda with signal noise. In this case, the noise in the signal can readily be calculated in this

---

[7] If N is zero, then the illiquidity measure in Equation (2) will also be zero, which is not reasonable. To get around this, we use a variable *ANA*, which is one plus the number of analysts. In this way, we avoid a sample bias because firms not covered by analysts are included in our sample (our approach mimics that of Brennan and Subrahmanyam (1995)).

[8] Note that since uninformed trades (z) follow the normal distribution, i.e., $z \sim N(0, v_z)$, $E[|z|] = \sqrt{\frac{2}{\pi}}std(z)$. Thus, $\sqrt{v_z} = std(z) = \sqrt{\frac{\pi}{2}}E[|z|]$, which, in turn, can be proxied by average trading volume.

context by considering the discrepancy between actual earnings and analysts' earnings forecasts. Thus, to estimate *ILLIQ_2* for each stock in each month, our proxy variables for each of the original variables in Equation (4) are as follows (as mentioned earlier, we consider alternative proxies for the inputs in Subsection 4.3):

P: The stock price at the previous month's end.

v_δ: This variable is proxied by EVOLA-sqr, which is the squared value of earnings volatility (EVOLA), where EVOLA is the standard deviation of earnings per share (EPS) from the most recent eight quarters.

v_ε: This variable is proxied by ESURP-sqr, which is the squared value of the earnings surprise (ESURP) defined as the absolute value of the current EPS minus the EPS forecast four quarters ago, scaled by the standard deviation of the most recent eight quarterly earnings.

v_z: We proxy this variable by AVG(DVOL)-sqr, which is the squared value of the average of daily dollar volume (in million dollars) within the previous month. To obtain this variable for each month, we use firms that have at least ten daily trading records in the previous month from the CRSP daily file.

1.4 Estimation of the illiquidity measures

Since volume data for NASDAQ stocks are not available prior to 1983, we first present the results for NYSE/AMEX (interchangeably, the "exchange market") stocks. Results for NASDAQ (interchangeably, the "OTC market") stocks are discussed in Subsection 4.3. To estimate our illiquidity measures (*ILLIQ_1* and *ILLIQ_2*) according to Equation (2) and Equation (4), the input variables related to the number of analysts (*ANA*), earnings surprise (*ESURP*, *ESURP-sqr*), and earnings volatility (*EVOLA*, *EVOLA-sqr*) are extracted from the I/B/E/S database. If a firm has one or more missing value(s) in the number of analysts, the missing months are filled with the previous month's value up to two quarters. We use the CRSP daily and monthly files to obtain other input variables: *STD*(*RET*), *AVG*(*DVOL*), *AVG*(*DVOL*)-*sqr*, and P. The average numbers of component stocks used each month to estimate *ILLIQ_1* and *ILLIQ_2* for NYSE/AMEX stocks are 1826.2 and 1696.7, respectively, over the past 324 months from January 1976 to December 2002.

Table 1 contains the descriptive statistics of the input variables employed to estimate our illiquidity measures. For *ILLIQ_1*, the average number of analysts is 6.19, the average standard deviation of daily returns is 2.7%, and the average of daily dollar volume is six million for NYSE/AMEX stocks over the sample period. In the case of *ILLIQ_2*, the average price level in the exchange market is $33.15, and the mean values of EVOLA-sqr, ESURP-sqr, and AVG(DVOL)-sqr are 199.53, 79.97, and 1072.25, respectively. We see that the input variables for the second measure are skewed to the left.

Having described our estimation procedure, we discuss and reiterate some advantages of our approach relative to estimation of illiquidity using intradaily data (e.g., as in Brennan and Subrahmanyam, 1996 or Sadka, 2006). In addition

Table 1
Descriptive statistics of the input variables in the two illiquidity measures for NYSE/AMEX stocks

Original input variables	Proxy variables employed	For *ILLIQ_1*			For *ILLIQ_2*		
		Mean	Median	STD	Mean	Median	STD
N	*ANA*	6.19	2.68	7.28	6.43	2.92	7.43
std(R)	*STD(RET)*	0.027	0.021	0.023	–	–	–
std(z)	*AVG(DVOL)*	6.00	0.60	19.29	–	–	–
P	P	–	–	–	33.15	18.76	433.23
ν_δ	*EVOLA-sqr*	–	–	–	199.53	0.06	7231.39
ν_ϵ	*ESURP-sqr*	–	–	–	79.97	0.02	3352.48
ν_z	*AVG(DVOL)-sqr*	–	–	–	1072.25	0.86	12699.00

This table reports descriptive statistics of the input variables of our two theoretically derived illiquidity measures, $ILLIQ\_1 = \frac{N^{0.5} std(R)}{(N+1) std(z)}$, and $ILLIQ\_2 = P^{-1} \frac{\nu_\delta}{(N+1)\nu_\delta + 2\nu_\epsilon} \sqrt{\frac{N(\nu_\delta+\nu_\epsilon)}{\nu_z}}$, where each input variable is defined as follows: N: the number of informed traders; *std(R)*: standard deviation of returns; *std(z)*: standard deviation of noise trades; P: asset price; ν_δ: variance of payoff innovations; ν_ϵ: variance of signal innovations; and ν_z: variance of noise trades. The above original input variables are in turn proxied by the variables shown in the second column of the table below. Each proxy variable is defined as follows: *ANA*: one plus the number of analysts following a firm; *STD(RET)*: standard deviation of daily returns in the previous month; *AVG(DVOL)*: average of daily dollar volume (in million dollars) in the previous month; *P*: month-end stock price of the previous month; *EVOLA-sqr*: squared value of earnings volatility (*EVOLA*), which is defined as standard deviation of EPSs from the most recent eight quarters; *ESURP-sqr*: squared value of earnings surprise (*ESURP*), which is defined as the absolute value of the current earnings per share (EPS) minus the EPS forecast four quarters ago; and *AVG(DVOL)-sqr*: squared value of *AVG(DVOL)*. The sample period is the past 324 months (twenty-seven years: 197601–200212) for NYSE/AMEX 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 numbers of component stocks used each month for *ILLIQ_1* and *ILLIQ_2* are 1826.2 and 1696.7, respectively.

to using closed-form expressions obtained from theory, our estimation method also avoids microstructural issues, such as price discreteness, inventory concerns, and the appropriate aggregation interval for order flows. Furthermore, unlike estimates based on regressions involving transactions data, thin trading does not affect the reliability of our estimates. Last, our method enables us to use a far broader cross-section and longer time series of data because we do not need to process the Institute for the Study of Securities Markets (ISSM) and TAQ databases, which are not available prior to 1983. We realize that despite all these advantages, the key challenge is to show whether our measures are priced after accounting for other popular illiquidity proxies; we will show below that this is indeed the case.

2. Methodology

For asset-pricing tests, we adopt the Brennan, Chordia, and Subrahmanyam (1998) approach. Assume that returns are generated by an L-factor approximate factor model

$$\tilde{R}_{jt} = E(\tilde{R}_{jt}) + \sum_{k=1}^{L} \beta_{jk} \tilde{f}_{kt} + \tilde{e}_{jt}, \tag{5}$$

where $\tilde{R}_{jt}$ is the return on security j at time t, and $\tilde{f}_{kt}$ is the unanticipated return on the kth factor ($k = 1, 2, \ldots, L$) 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

$$E(\tilde{R}_{jt}) - R_{Ft} = \sum_{k=1}^{L} \theta_{kt}\beta_{jk}, \tag{6}$$

where R_{Ft} is the return on the risk-free asset, and θ_{kt} is the risk premium on the factor portfolio k. Plugging Equation (6) into Equation (5), the APT implies that realized returns are given by

$$\tilde{R}_{jt} - R_{Ft} = \sum_{k=1}^{L} \beta_{jk}\tilde{F}_{kt} + \tilde{e}_{jt}, \tag{7}$$

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

Our goal is to test whether the two illiquidity measures derived in Section 1, based on the strategic microstructure model, have incremental explanatory power for returns relative to the Fama and French (1993) (henceforth, FF) three-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:

$$\tilde{R}_{jt+1} - R_{Ft+1} = c_0 + \phi\Gamma\_i_{jt} + \sum_{k=1}^{L} \theta_k\beta_{jkt} + \sum_{m=1}^{M} c_m Z_{mjt} + \tilde{e}_{jt+1}, \tag{8}$$

where Γ_i_{jt} ($i = 1$ or 2) is one of our illiquidity measures (*ILLIQ_1* or *ILLIQ_2*) for security j in month t estimated in Section 1, and a vector of control variables, Z_{mjt}, is firm characteristic m ($m = 1, \ldots, M$) for security j in month t. Under the null hypothesis that expected excess returns depend only on risk measured by β_{jk}, coefficients ϕ and c_m ($m = 1, \ldots, M$) 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 nonrisk 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.

To address the errors-in-variables problem, we use risk-adjusted returns as the dependent variables. Risk adjustment is done using the Fama–French (1993) factors (MKT_t, SMB_t, and HML_t) in two different ways. In the first method, we compute risk-adjusted returns, $\tilde{R}^{*1}_{jt}$, for each month as the sum of the intercept and the residual, i.e.,

$$\begin{aligned}\tilde{R}^*_{jt} &= (\tilde{R}_{jt} - R_{Ft}) - \left(\widehat{\beta}^*_{j1} MKT_t + \widehat{\beta}^*_{j2} SMB_t + \widehat{\beta}^*_{j3} HML_t\right) \\ &= \widehat{\alpha}^* + \widehat{\tilde{e}}^*_{jt},\end{aligned} \tag{9}$$

after conducting regressions in Equation (7) (but with a constant term α) using the *entire* sample range (from January 1976 to December 2002 for NYSE/AMEX stocks) of the data.[9] We call this risk-adjusted return ($\tilde{R}^*_{jt}$) *FF3-adj EXSRET1*. We also use another version of risk adjustment for robustness. Thus, 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 sixty months (at least twenty-four months) in Equation (7). Given the current month's data ($\tilde{R}_{jt}$-R_{Ft}, MKT_t, SMB_t, and HML_t) and the factor loadings ($\widehat{\beta}^{**}_{jk}$) 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:

$$\tilde{R}^*_{jt} = (\tilde{R}_{jt} - R_{Ft}) - \left(\widehat{\beta}^{**}_{j1} MKT_t + \widehat{\beta}^{**}_{j2} SMB_t + \widehat{\beta}^{**}_{j3} HML_t\right). \tag{10}$$

We call this risk-adjusted return ($\tilde{R}^*_{jt}$) *FF3-adj EXSRET2*.

The risk-adjusted returns from Equation (9) and Equation (10) constitute the raw material for the estimates that we present in the following Fama–MacBeth (1973) cross-sectional regressions:

$$\tilde{R}^*_{jt+1} = c_{0t} + \phi_t \Gamma\_i_{jt} + \sum_{m=1}^{M} c_{mt} Z_{mjt} + \tilde{e}'_{jt+1}, \quad i = 1, 2. \tag{11}$$

Note that the error term in Equation (11) is different from that in Equation (8) because the error in Equation (11) 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 Equation (11): the statistics based on regressions with the dependent variable in Equation (11) being (i) risk-*unadjusted* returns in excess of the one-month T-bill rate (we call this unadjusted return *EXSRET*); (ii) risk-adjusted excess returns using the first method, *FF3-adj EXSRET1*; and (iii) risk-adjusted excess returns using the second method, *FF3-adj EXSRET2*. For our purposes, we estimate the vector of coefficients $\mathbf{c}_t = [c_{0t}\ \phi_t\ c_{1t}\ c_{2t}\ \ldots\ c_{Mt}]'$ from Equation (11) each month with a simple OLS regression as

$$\widehat{\mathbf{c}}_t = (\mathbf{Z}_t'\mathbf{Z}_t)^{-1}\mathbf{Z}_t'\tilde{\mathbf{R}}^{*h}_{t+1},$$

where $h = 1$ or 2, $\mathbf{Z}_t = [\Gamma\_i \quad Z_1\ Z_2 ... Z_M]'$ and $\tilde{\mathbf{R}}^{*h}_{t+1}$ is the vector of risk-adjusted excess returns based on Equation (9) or Equation (10). The standard

[9] 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.

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 errors in Equation (7), these errors affect only the dependent variable, $\tilde{R}^{*h}_{t+1}$, as we see in Equations (9)–(11). While the factor loadings will be correlated with vector $\mathbf{Z}_t = [\Gamma\_i \;\; Z_1 \; Z_2 \; \ldots \; Z_M]'$, there is no *a priori* reason to believe that the errors in the estimated loadings will be correlated with vector $\mathbf{Z}_t$. This implies that coefficient vector $\widehat{\mathbf{c}}_t$ estimated in Equation (11) is unbiased.

If the errors in the estimated factor loadings are correlated with 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 *FF3-adj EXSRET1* and *FF3-adj EXSRET2*: i.e., the constant term (and its t-value) from the regression of the monthly coefficients ($\widehat{\mathbf{c}}_t$) estimated in Equation (11) on the time series of FF three-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 $\mathbf{Z}_t$.

3. Data, Definitions, Descriptive Statistics, and Adjustments

For examining the cross-section of expected returns, we mainly use the NYSE/AMEX data at a daily and/or monthly frequency over the 324 months (twenty-seven years: 197601–200212). In those cases in which accounting variables and other data are available, only on a yearly (or quarterly) basis, we keep the relevant values constant for twelve months (or three months) in the regressions.[10]

The three dependent variables (*EXSRET*, *FF3-adj EXSRET1*, and *FF3-adj EXSRET2*) defined in Section 2 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 (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/). In addition to the variables mentioned above, we use six firm-specific characteristics in the regressions as control variables: *SIZE*, *BTM*, *MOM1–MOM4*. The definitions of the control and related variables are as follows:

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

[10] The data series available only on a yearly basis are the variables related to the book-to-market ratio (*BM_Raw*, *BM_Trim*, and *BTM*) and the effective cost measure (*Roll_Gibbs*: to be explained later). Those available only on a quarterly basis are accounting performance-related variables, ESURP-sqr, and EVOLA-sqr.

SIZE: The natural logarithm of *MV*.

BM_Raw: The book-to-market ratio defined as *BV/MV*, where the book value (BV) is common equity plus deferred taxes (in million dollars).

BM_Trim: The Winsorized book-to-market ratio, where *BM_Raw* values greater than the 99.5 percentile value or less than the 0.5 percentile value in a month are set equal to the 99.5 and 0.5 percentile values, respectively.

BTM: The natural logarithm of *BM_Trim*. Following Fama and French (1992), we fill monthly *BM_Raw* (hence *BM_Trim* and *BTM*) values for July of year t to June of year $t + 1$ with the value computed using the accounting data at the end of year $t - 1$, 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 three months (from month $t - 1$ to month $t - 3$).

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

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

MOM4: The compounded holding-period return over the three 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 5 we run a horse race to compare the effects of our two illiquidity measures with those of three other alternative (il)liquidity measures commonly used in the literature. The alternative measures to be analyzed in our study are notated and defined as follows:

Amihud: The illiquidity measure of Amihud (2002). We estimate this measure each month as the average of $|r|/DDVOL$, where r is the daily stock return and *DDVOL* is the daily dollar volume in thousands.

Roll_Gibbs: The 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 (http://www.stern.nyu.edu/~jhasbrou).

DVOL: The average of daily dollar volume (in thousand dollars) within each month for each stock.

TURN: The average of daily share turnover values within each month for each stock.

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 three (il)liquidity measures (*Amihud*, *DVOL*, and *TURN*) are estimated using the CRSP daily and monthly files. The average number of component stocks used each month in the Fama–MacBeth (1973) cross-sectional regressions for NYSE/AMEX stocks is 1597–1715.

Table 2 reports the time-series average values of monthly means, medians, standard deviations (STD), and other descriptive statistics for the key variables. The values of each statistic are first computed cross-sectionally and then averaged in the time series over the sample period. A noteworthy aspect is that *ILLIQ_1* and *ILLIQ_2* are highly leptokurtic, as well as significantly skewed to the left. The large kurtoses and skewnesses of *ILLIQ_1* and *ILLIQ_2* also imply that sample distributions of the two measures exhibit many extreme observations. To alleviate this problem, Hasbrouck (1999, 2005, 2006) advocates employing the square-root transform of liquidity measures.[11] Following his suggestion, we also apply square-root transformation to the two raw measures. As we see in Table 2, the skewnesses and kurtoses of the corresponding measures ($[ILLIQ\_1]^{1/2}$, $[ILLIQ\_2]^{1/2}$) are substantially reduced by the transformation. For this reason, we perform our analyses mainly using the square-root transforms of the two theory-based measures, rather than using the raw measures themselves.[12]

To examine the time-series behavior of our two illiquidity measures, we plot in Figure 1 the value-weighted series of the transformed measures over the sample period. As we see in Figure 1(a), value-weighted $[ILLIQ\_1]^{1/2}$ of NYSE/AMEX stocks demonstrates a very high level in the initial part of the sample period, possibly due to the aftermath of the oil crisis in late 1973. However, it generally exhibits a decreasing time trend, suggesting that market liquidity has improved since the mid-1970s. As can be seen in Figure 1(b), the trend of $[ILLIQ\_2]^{1/2}$ is qualitatively similar to that of $[ILLIQ\_1]^{1/2}$, with its volatility being more pronounced. For brevity, we do not report the graphs based on the series of the equal-weighted illiquidity measures. But their trends are qualitatively similar, with the absolute levels generally being higher than those of the value-weighted series.

Table 2 also shows that the average market value (*MV*) is \$1.93 billion, and the book-to-market ratio (BM_Raw) is 0.98 on average for NYSE/AMEX stocks over the sample period. Both variables tend to be left-skewed.

Next, we examine the average correlation coefficients between our explanatory variables in Table 3.[13] Our two illiquidity measures are highly correlated, with coefficients of 67% between the two untransformed measures and 80% between the two transformed measures. The two measures are negatively correlated with the four momentum variables, suggesting that good past price performance of a stock tends to contribute to the improvement in the liquidity

[11] Hasbrouck (2005) eschews the logarithmic transformation because it is theoretically possible for illiquidity to be zero. Given that our two illiquidity measures often have a value of zero, we opt to use the square-root instead of the log transform.

[12] With the square-root transformation, however, Table 2 shows that our two transformed measures are still highly skewed. This raises a question that our empirical results that follow may be driven purely by a few outlier stocks. This issue will be examined further in Sections 4 and 5.

[13] 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%).

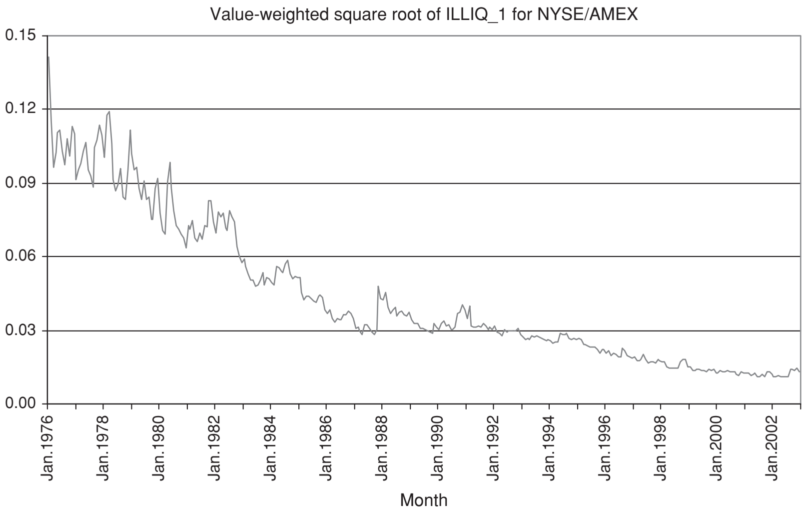
Table 2
Descriptive statistics of the key variables for NYSE/AMEX stocks

Variables	Mean	Median	STD	CV	Skewness	Kurtosis
ILLIQ_1	1.55	0.04	13.76	734.57	18.22	477.94
[*ILLIQ_1*]$^{1/2}$	0.49	0.18	1.01	210.02	6.52	82.30
ILLIQ_2	8.93	0.02	167.11	1295.59	24.65	750.37
[*ILLIQJ*]$^{1/2}$	0.65	0.13	2.32	346.98	11.50	217.31
MV	1931.69	275.42	6830.84	352.44	11.61	203.15
SIZE	5.35	5.35	2.00	37.67	0.04	−0.45
BM_Raw	0.98	0.84	0.76	78.59	4.93	80.92
BM_Trim	0.98	0.84	0.68	70.27	2.18	9.13
BTM	−0.30	−0.22	0.71	−443.26	−0.83	2.91
MOM1	0.016	0.021	0.192	352.77	−0.32	6.02
MOM2	0.017	0.021	0.188	−1835.49	−0.22	5.33
MOM3	0.020	0.023	0.185	294.54	−0.12	4.73
MOM4	0.022	0.025	0.184	40.77	−0.06	4.40

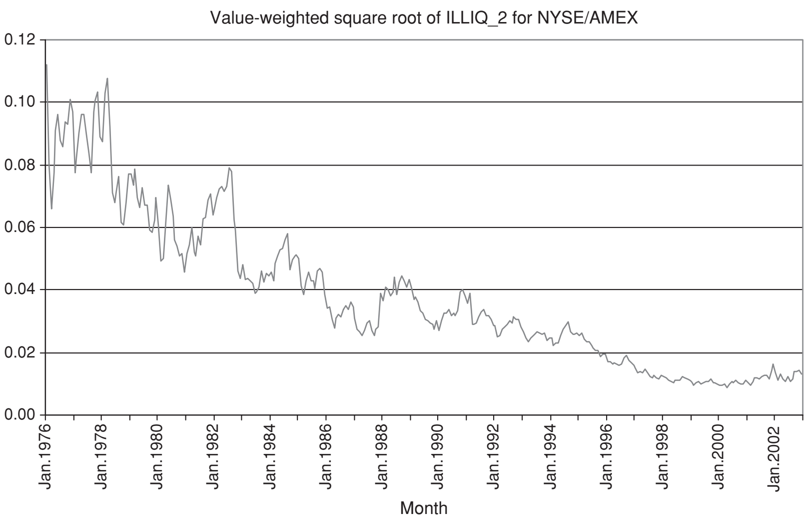
This table reports descriptive statistics (mean, median, standard deviation (STD), coefficient of variation (CV), skewness, and kurtosis) of the key variables to be used on the right-hand side in the Fama–MacBeth (1973) cross-sectional regressions. Each variable is defined as follows: *ILLIQ_1*: the first illiquidity measure defined as in Table 1; *ILLIQ_2*: the second illiquidity measure defined as in Table 1; *MV*: market value defined as the month-end stock price times the number of shares outstanding (in million dollars); *SIZE*: natural logarithm of *MV*; *BM_Raw*: the untrimmed book-to-market ratio defined as *BV/MV*, where the book value (*BV*) is common equity plus deferred taxes (in million dollars); *BM_Trim*: the Winsorized book-to-market ratio, where *BM_Raw* values greater than the 99.5 percentile value or less than the 0.5 percentile value in a month are set equal to the 99.5 and 0.5 percentile values, respectively; *BTM*: natural logarithm of *BM_Trim*; *MOM1*: compounded holding-period return of a stock over the most recent three months (from month $t-1$ to month $t-3$); *MOM2*: compounded holding-period return over the next recent three months (from month $t-4$ to month $t-6$); *MOM3*: compounded holding-period return over the three months from month $t-7$ to month $t-9$; *MOM4*: compounded holding-period return over the three months from month $t-9$ to month $t-12$. The sample period is of the past 324 months (twenty-seven years: 197601–200212) for NYSE/AMEX 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 1826 (except that it is 1697 for *ILLIQ_2*).

of that stock. It is also not surprising to observe that correlation of SIZE with the two illiquidity measures is negative and statistically significant at any conventional level, because we would expect larger firms with greater breadth of ownership to be more liquid than smaller ones. The coefficients of correlation between the book-to-market ratio and the two illiquidity measures are positive and statistically significant. This indicates that value stocks are likely to be more illiquid.

Some of our time series may be nonstationary. 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 nonstationarity are our illiquidity measures and their square-root transforms (*ILLIQ_1*, *ILLIQ_2*; [*ILLIQ_1*]$^{1/2}$, [*ILLIQ_2*]$^{1/2}$), the alternative (il)liquidity measures and their square-root transforms (*Amihud*, *Roll_Gibbs*, *DVOL*, *TURN*; [*Amihud*]$^{1/2}$, [*Roll_Gibbs*]$^{1/2}$, [*DVOL*]$^{1/2}$, [*TURN*]$^{1/2}$), and *SIZE*. To eliminate nonstationarity, we adjust these data series in two steps along the lines of Gallant, Rossi, and Tauchen (1992) before conducting cross-sectional regressions in the next sections (Chordia,



(a)



(b)

Figure 1
Trends in the value-weighted illiquidity measures for NYSE/AMEX stocks
The graphs show the trends of the market value-weighted series of $[ILLIQ\_1]^{1/2}$ and $[ILLIQ\_2]^{1/2}$ for NYSE/AMEX stocks over the past 324 months (twenty-seven years: 197601–200212). Each of the two graphs is a time-series plot of the monthly cross-sectional (market-value weighted) averages over the sample period. *ILLIQ_1* and $ILLIQ\_2$ are defined as in Table 1. Figure 1(a) is for $[ILLIQ\_1]^{1/2}$, and Figure 1(b) is for $[ILLIQ\_2]^{1/2}$. The average numbers of component stocks used each month are 1826 for $[ILLIQ\_1]^{1/2}$ and 1697 for $[ILLIQ\_2]^{1/2}$.

Table 3
Correlations between explanatory variables for NYSE/AMEX stocks

	ILLIQ_1	[*ILLIQ_1*]$^{1/2}$	*ILLIQ_2*	[*ILLIQ_2*]$^{1/2}$	*SIZE*	*BTM*	*MOM1*	*MOM2*	*MOM3*	*MOM4*
ILLIQ_1	1									
[ILLIQ_1]$^{1/2}$	0.849	1								
ILLIQ_2	0.673	0.546	1							
[ILLIQ_2]$^{1/2}$	0.746	0.798	0.855	1						
SIZE	−0.301	−0.627	−0.186	−0.455	1					
BTM	0.090	0.182	0.060	0.157	−0.290	1				
MOM1	−0.090	−0.141	−0.076	−0.147	0.122	0.043	1			
MOM2	−0.080	−0.131	−0.060	−0.125	0.120	0.039	0.015	1		
MOM3	−0.075	−0.126	−0.058	−0.123	0.121	0.009	0.047	0.012	1	
MON4	−0.076	−0.125	−0.056	−0.122	0.122	−0.034	0.052	0.044	0.006	1

This table shows the average correlations between the key variables for NYSE/AMEX stocks over the 324 months (twenty-seven years: 197601–200212). 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: *ILLIQ_1*: the first illiquidity measure defined as in Table 1; *ILLIQ_2*: the second illiquidity measure defined as in Table 2; *SIZE*: natural logarithm of *MV*, which is defined as the month-end stock price times the number of shares outstanding (in million dollars); *BTM*: natural logarithm of *BM_Trim*, which is the Winsorized book-to-market ratio, where book-to-market ratios greater than the 99.5 percentile value or less than the 0.5 percentile value in a month are set equal to the 99.5 and 0.5 percentile values, respectively; *MOM1*: compounded holding-period return of a stock over the most recent three months (from month $t-1$ to month $t-3$); *MOM2*: compounded holding-period return over the next recent three months (from month $t-4$ to month $t-6$); *MOM3*: compounded holding-period return over the three months from month $t-7$ to month $t-9$; *MOM4*: compounded holding-period return over the three months from month $t-9$ to month $t-12$. The average number of component stocks used in a month is 1826 (except that it is 1697 for *ILLIQ_2*).

Huh, and Subrahmanyam (2007) use a similar technique). 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 following mean equation:

$$\kappa = x'\psi + \xi, \tag{12}$$

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

$$\log(\xi^2) = x'\theta + \epsilon. \tag{13}$$

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

$$\kappa_{adj} = a + b\{\widehat{\xi}/\exp(x'\theta/2)\}, \tag{14}$$

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 (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). Thus, henceforth we use the adjusted values of size and the illiquidity variables.

4. Empirical Results

4.1 Features of the portfolios formed on illiquidity and size

Before moving on to regression analyses, we report mean returns for the twenty-five portfolios formed by sorting the component stocks into quintiles by illiquidity (*ILLIQ_1*) and size (market capitalization, *MV*), as defined in the previous sections.[14] The size and illiquidity groups are notated as *Size i* and *Illiq i* ($i = 1$–5), respectively. For consistency in our empirical results, the component stocks used for the portfolio analyses are limited to those used for our main regression analyses to follow.[15] Furthermore, to reduce a possible upward bias in computing portfolio returns due to the exclusion of delisting returns (Shumway, 1997), we include all relevant delisting returns over the sample period.[16] To minimize another possible bias in equal-weighted portfolio returns documented by Blume and Stambaugh (1983), we compute value-weighted portfolio returns (using market capitalization as of the end of the previous month as weights). The time-series means of these returns appear in Table 4. The average number of component stocks used in each of the 25 portfolios in each month is 68.61. The middle part of Table 4 (row of high–low) shows the time-series averages of the return differences between the highest and lowest illiquidity portfolios within a given size group. Each of the portfolio returns is followed by a t-statistic for testing the null hypothesis that the time-series average of the return differences equals zero.

[14] While we report results for sequential sorts, independently sorting by size and illiquidity leads to similar results, which are available from the authors. Also, while we present results for value-weighted portfolios, equally weighting returns does not materially change our conclusions.

[15] These are the stocks that have all the required variables for regression specification *EXSRET* in Panel A of Table 6, as well as for regression specifications 1 and 9 in Table 9.

[16] However, we find that the effect of the upward bias in portfolio returns by excluding delisting returns is small. One reason for this is that the number of delisted firms used in our study is relatively small. Second, the sample in our study mainly consists of stocks listed on (or delisted from) the NYSE/AMEX, for which the delisting bias is not as severe as for NASDAQ stocks. This aspect is pointed out in Shumway (1997). Third, if firms were delisted for performance-related reasons (delisting code (DLSTCD) 500-591 in CRSP), while their delisting returns are usually negative, they are also often positive if the firms were delisted because “prices fell below the acceptable level” (e.g., delisting code 552). Last, delisting returns are also mostly positive if firms were delisted because they were merged with other firms (delisting code 200-290).

Table 4
Value-weighted average returns (including delisting returns) for portfolios formed on size and illiquidity with NYSE/AMEX stocks

Value-weighted average returns (including delisting returns) for 25 portfolios

	size group				
Illiquidity group	Size 1 (small)	Size 2	Size 3	Size 4	Size 5 (big)
Illiq 1 (low)	0.0183	0.0151	0.0122	0.0113	0.0075
	5.11	*5.01*	*4.36*	*4.39*	*3.10*
Illiq 2	0.0229	0.0182	0.0140	0.0131	0.0091
	7.30	*6.43*	*4.95*	*5.21*	*3.70*
Illiq 3	0.0279	0.0200	0.0183	0.0143	0.0091
	8.03	*6.28*	*6.15*	*5.21*	*3.47*
Illiq 4	0.0294	0.0231	0.0165	0.0156	0.0071
	7.92	*6.84*	*5.49*	*5.48*	*2.47*
Illiq 5 (high)	0.0374	0.0282	0.0217	0.0170	0.0084
	7.92	*7.65*	*6.44*	*5.29*	*2.79*
High–low	0.0191	0.0131	0.0094	0.0057	0.0010
	5.59	*5.08*	*4.03*	*2.64*	*0.46*

This table reports time-series averages of monthly value-weighted cross-sectional mean returns for portfolios obtained by sorting component stocks sequentially into quintiles by illiquidity and market capitalization (*ILLIQ_1* and *MV*, as defined in Section 3). Size and illiquidity groups are notated as *size i* and *illiq i* (i = 1–5), respectively. The upper part of the table contains the time-series average of value-weighted cross-sectional mean returns for each of the twenty-five portfolios, together with their t-statistics (italicized). In computing the twenty-five portfolio returns, delisting returns (*DLRET* in CRSP) from seventy-three delisted firms are included over the sample period. The lower part of the table shows the time-series average returns for the five differential portfolios (each of which contains a time series of return differences between the highest illiquidity portfolio (high) and the lowest illiquidity portfolio (low) within a given 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 324 months (twenty-seven years: 197601–200212) for NYSE/AMEX stocks. The average number of component stocks in each portfolio in a month is 68.61.

Table 4 shows that within a given size group, the value-weighted average return increases with illiquidity, except for the largest size group. In particular, the row of high–low demonstrates that the average returns are positive and significantly different from zero in four out of the five cases. This finding suggests that illiquidity pricing is stronger in smaller firms, though not solely driven by a few small firms. This aspect is re-examined in later analyses (viz., Table 10).

To get additional perspective on the impact of illiquidity on returns, we report in Table 5 the intercepts and t-statistics from the time-series regressions of the portfolio returns (in excess of the one-month T-bill rate) on the FF factors. The table demonstrates that the value-weighted portfolio returns over and above the returns predicted by the FF three-factor model (denoted by *FF3 alphas*) are positive and significantly different from zero at the 5% level for portfolios consisting of smaller (*Size* 1–*Size 3*) and more illiquid (*Illiq 2*–*Illiq 5*) stocks. However, the abnormal returns for portfolios with bigger and/or very liquid stocks are not always positive.

In the middle part of Table 5, we provide the intercepts from the time-series regressions of the return differences between the most and least illiquid portfolios within a given size group on the FF factors, together with the associated t-statistics. The table indicates that the abnormal returns for four out of the five differential (*high–low*) portfolios are positive and three of them are significantly

Table 5
Intercepts (FF3 alphas) from the time-series regressions of excess returns (including delisting returns) on the Fama–French three factors and GRS tests for portfolios formed on size and illiquidity with NYSE/AMEX stocks

FF3 alphas from the time series of value-weighted returns

	Size group				
Illiquidity group	Size 1 (small)	Size 2	Size 3	Size 4	Size 5 (big)
Illiq 1 (low)	0.0030	0.0009	−0.0007	−0.0004	−0.0019
	1.46	*0.64*	*−0.63*	*−0.42*	*−2.50*
Illiq 2	0.0084	0.0044	0.0003	0.0009	−0.0009
	5.11	*3.30*	*0.24*	*0.81*	*−1.17*
Illiq 3	0.0119	0.0051	0.0037	0.0007	−0.0026
	6.28	*3.03*	*2.52*	*0.51*	*−2.51*
Illiq 4	0.0140	0.0080	0.0018	0.0013	−0.0056
	5.89	*4.21*	*1.30*	*0.90*	*−3.85*
Illiq 5 (high)	0.0211	0.0123	0.0061	0.0020	−0.0051
	6.28	*5.13*	*2.83*	*1.08*	*−3.32*
High–low	0.0182	0.0113	0.0068	0.0024	−0.0032
	5.36	*4.58*	*3.41*	*1.29*	*−1.94*

GRS test for five (high–low) portfolios

Null hypothesis	F-statistic	p-value
H_0^{Diff}: FF3 alphas for five (high–low) portfolios are jointly 0	*1.82*	*0.0110*

This table reports the regression intercepts (FF3 alphas) and the GRS test result for the portfolios formed by sorting stocks sequentially into quintiles by illiquidity and market capitalization (*ILLIQ_1* and *MV*, as defined in Section 3). Size and illiquidity groups are notated as *size i* and *illiq i* (i = 1–5), respectively. The twenty-five FF3 alphas contained in the upper part of the table are the intercepts from the time-series regressions of the value-weighted portfolio returns (in excess of the one-month T-bill rate) on the Fama–French three factors. In computing the twenty-five portfolio returns, delisting returns (*DLRET* in CRSP) from seventy-three delisted firms are included over the sample period. The values italicized in the second row of each illiquidity group in the upper part are t-statistics from the time-series regressions. The five FF3 alphas contained in the middle part of the table are the intercepts from the time-series regressions of the return difference (between the highest illiquidity portfolio (high) and the lowest illiquidity portfolio (low) within a given size group) on the Fama–French three factors, together with their t-statistics from the regressions. The lower part of the table reports the GRS test result for the null hypothesis (H_0^{Diff}) that the FF3 alphas for the five (high–low) 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 324 months (twenty-seven years: 197601–200212) for NYSE/AMEX stocks. The average number of component stocks in each portfolio in a month is 68.61.

different from zero at the 5% level. These results are broadly consistent with those in the last row of Table 4.

Following Fama and French (1993), we test the hypothesis that the intercepts for the five (*high–low*) portfolios are jointly zero, using the Gibbons, Ross, and Shanken (1989) statistic. This F-statistic and its corresponding p-value are listed at the bottom of Table 5, and provide evidence against the null hypothesis at the 5% level of significance. Of course, the portfolio analysis is preliminary in a sense that it does not account for other characteristics that may affect stock returns. Moreover, portfolio returns averaged across stocks may conceal interesting dynamics 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.

4.2 Cross-sectional regressions

We have observed in Table 4 (Table 5) that within a given size group, the average portfolio return (the portfolio abnormal return) is likely to increase with illiquidity, suggesting that theory-based illiquidity is a priced factor. In this section, we formally test whether our two illiquidity measures have any impact on returns. As mentioned above, our test involves the following cross-sectional regression estimated at the monthly frequency:

$$\tilde{R}^{*}_{jt+1} = c_{0t} + \phi_t \Gamma\_i_{jt} + \sum_{m=1}^{M} c_{mt} Z_{mjt} + \tilde{e}'_{jt+1}, \quad i = 1, 2, \tag{15}$$

where $\tilde{R}^{*}_{jt+1}$ represents either the risk-unadjusted excess return (*EXSRET*) or one of the two risk-adjusted excess returns (*FF3-adj EXSRET1* and *FF3-adj EXSRET2*) defined and estimated in Section 2, Γ_i_{jt} is either of our two theory-based illiquidity measures (*ILLIQ_1*, *ILLIQ_2*, and their square-root transforms) derived and estimated in Section 1, and Z_{mjt} denotes firm characteristic m for stock j in month t.[17]

We conduct cross-sectional regressions based on Equation (15) using both the two raw illiquidity measures (*ILLIQ_1*, *ILLIQ_2*) and their square-root transforms ($[ILLIQ\_1]^{1/2}$, $[ILLIQ\_2]^{1/2}$). To conserve space, however, we report the results only with the square-root transformed measures in the analyses throughout the paper.[18] The standard Fama–MacBeth statistics (the time-series average of the estimated coefficients from the equation above and its t-statistic) are presented in Table 6.[19] Along with the average coefficients and t-statistics, we also list the average 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.*). As mentioned in Section 3, for some nonstationary variables we use the GRT-adjusted series (indicated by superscript "a" in the tables in this and the next section) for the cross-sectional regressions.

As we see in Panel A of Table 6, the average number of component stocks used in the monthly regressions with our first illiquidity measure, $[ILLIQ\_1]^{1/2}$,

[17] Avramov and Chordia (2006) show that a constant-beta version of the Fama and French (1993) three-factor model cannot adequately capture the predictive ability of firm characteristics in stock returns. Thus, we control for characteristics (Z_{mjt}) such as size, book-to-market equity, and past returns when we examine the impact of our illiquidity measures on FF3-adjusted returns in Equation (15).

[18] The reason has already been explained in Section 3 (see Table 2 and its related text). However, the results with the raw illiquidity measures (*ILLIQ_1* and *ILLIQ_2*) are also available upon request from the authors.

[19] 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%). 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 6, which includes both types of t-statistics, shows that the levels of t-values are very close 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.

Table 6
Results of monthly cross-sectional regressions with [*ILLIQ_1*]$^{1/2}$ and [*ILLIQ_2*]$^{1/2}$ for NYSE/AMEX stocks

Panel A: With square root of *ILLIQ_1*				Panel B: With square root of ILLIQ_2			
Expla. variables	*EXSRET*	*FF3-adj EXSRET1*	*FF3-adj EXSRET2*	Expla. variables	*EXSRET*	*FF3-adj EXSRET1*	*FF3-adj EXSRET2*
Intercept	1.895**	0.880**	1.006**	Intercept	2.103**	1.075**	1.210**
	5.51	*6.04*	*6.61*		*6.09*	*6.78*	*7.56*
[*ILLIQ_1*]$^{1/2,a}$	0.224**	0.204**	0.219**	[*ILLIQ_2*]$^{1/2,a}$	0.099**	0.089**	0.097**
	4.36	*4.15*	*4.16*		*3.92*	*3.64*	*3.82*
SIZE^a	−0.271**	−0.209**	−0.223**	*SIZE*^{a}	−0.295**	−0.231**	−0.247**
	−7.39	*−10.35*	*−10.44*		*−7.68*	*−10.32*	*−10.71*
BTM	0.079	0.041	−0.018	*BTM*	0.040	0.007	−0.049
	1.12	*0.88*	*−0.38*		*0.58*	*0.15*	*−1.04*
MOM1	0.669*	0.785**	0.848**	*MOM1*	0.635	0.737**	0.844**
	2.00	*2.86*	*2.74*		*1.88*	*2.67*	*2.77*
MOM2	1.348**	1.279**	1.316**	*MOM2*	1.298**	1.238**	1.333**
	4.82	*5.34*	*4.86*		*4.57*	*5.14*	*4.94*
MOM3	1.405**	1.303**	1.051**	*MOM3*	1.361**	1.257**	1.042**
	4.92	*5.18*	*3.61*		*4.74*	*4.94*	*3.54*
MOM4	1.089**	0.999**	0.776**	*MOM4*	1.012**	0.943**	0.703**
	4.65	*4.82*	*2.93*		*4.18*	*4.40*	*2.62*
Avg. *R*-sqr.	0.046	0.028	0.030	Avg. *R*-sqr.	0.048	0.030	0.031
Avg. obs.	1715.3	1715.3	1676.7	Avg. obs.	1607.6	1607.6	1596.5

This table reports the monthly Fama–MacBeth (1973)-type cross-sectional regressions using the square-root values of *ILLIQ_1* (i.e., [*ILLIQ_1*]$^{1/2}$) (in Panel A) and the square-root values of *ILLIQ_2* (i.e., [*ILLIQ_2*]$^{1/2}$) (in Panel B) for NYSE/AMEX stocks over the 324 months (twenty-seven years: 197601–200212). The dependent variables (*EXSRET*, *FF3-adj EXSRET1*, and *FF3-adj EXSRET2*) are all one-month leading. 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; *FF3-adj EXSRET1*: the risk-adjusted excess return using the Fama–French (FF) three 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; *FF3-adj EXSRET2*: the risk-adjusted excess return using the FF3 factors with factor loadings being estimated from the five-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 (α, $\beta_1\beta_2$, β_3) are first estimated for each month using the time-series data of the past sixty months in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \epsilon$, 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 FF3 factors; [*ILLIQ_1*]$^{1/2}$: square root of *ILLIQ_1*, which is defined as $\frac{N^{0.5} std(R)}{(N+1) std(z)}$, where *N* is the number of informed traders, *std(R)* is the standard deviation of returns, and *std(z)* is the standard deviation of noise trades (the original input variables are proxied by the variables as shown in Table 1); [*ILLIQ_2*]$^{1/2}$: square root of *ILLIQ_2*, which is defined as $P^{-1}\frac{\nu_\delta}{(N+1)\nu_\delta + 2\nu_\epsilon}\sqrt{\frac{N(\nu_\delta+\nu_\epsilon)}{\nu_z}}$, where *P* is the asset price, *N* is the number of informed traders, ν_δ is the variance of payoff innovations, ν_ϵ is the variance of signal innovations, and ν_z is the variance of noise trades (the original input variables are proxied by the variables as shown in Table 1); *SIZE*: natural logarithm of *MV*, which is defined as the month-end stock price times the number of shares outstanding (in million dollars); *BTM*: natural logarithm of *BM_Trim*, which is the Winsorized book-to-market ratio, where book-to-market ratios greater than the 99.5 percentile value or less than the 0.5 percentile value in a month are set equal to the 99.5 and 0.5 percentile values, respectively; *MOM1*: compounded holding-period return of a stock over the most recent three months (from month $t-1$ to month $t-3$); *MOM2*: compounded holding-period return over the next recent three months (from month $t-4$ to month $t-6$); *MOM3*: compounded holding-period return over the three months from month $t-7$ to month $t-9$; *MOM4*: compounded holding-period return over the three months from month $t-9$ to month $t-12$. To remove nonstationarity, the Gallant, Rossi, and Tauchen (GRT, 1993) procedure has been applied to [*ILLIQ_1*]$^{1/2}$, [*ILLIQ_2*]$^{1/2}$, and *SIZE* 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. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

for NYSE/AMEX stocks ranges from 1676.7 to 1715.3, depending on data availability of the variables. *Avg. R-sqr.* is about 2.8%–4.6%. The explanatory power of the regressions is higher with the unadjusted excess returns (*EXSRET*) than with the risk-adjusted returns (*FF3-adj EXSRET1*, *FF3-adj EXSRET2*). This suggests that the Fama–French model has some ability to price the cross-section of stocks. Given that our second illiquidity measure, $[ILLIQ\_2]^{1/2}$, requires more input variables, Panel B shows that the average number of component stocks used in the regressions with this illiquidity measure decreases to 1596.5–1607.6. Other aspects of the explanatory power are similar to those with $[ILLIQ\_1]^{1/2}$.

We first discuss the results from the Fama–MacBeth regressions of *EXSRET* on $[ILLIQ\_1]^{1/2}$, as well as other firm-level characteristics that are known to be associated with expected returns, namely, *SIZE*, *BTM*, and the four momentum variables (*MOM1–MOM4*). The second column of Panel A in Table 6 exhibits that the average coefficients of $[ILLIQ\_1]^{1/2}$ are positive and statistically significant at the 1% level after controlling for other firm-level characteristics, confirming the hypothesis that stocks with higher illiquidity are expected to have higher (excess) returns. Consistent with the prior literature, the coefficients of *SIZE* are negative and statistically significant. But the coefficient of *BTM* is not significant. The momentum variables are all strongly positively related to returns.

We now consider whether the relations observed above are maintained when the dependent variable is risk-adjusted using the FF factors. The estimates of illiquidity and characteristic rewards ($\widehat{\phi}$ and $\widehat{c}_m$'s) for returns, adjusted by the first method in Section 2 (*FF3-adj EXSRET1*), are presented in the next column of Panel A in Table 6. By risk-adjusting, some of the coefficients in the right-hand-side variables tend to attenuate, but the relations are essentially similar, with the levels of statistical significance becoming even reinforced in many cases. $[ILLIQ\_1]^{1/2}$ continues to be strongly positively related to risk-adjusted returns. With the risk-adjustment, firm size and *BTM* show similar patterns as above. Overall, *SIZE* plays an important role in predicting stock returns in the NYSE/AMEX market. The four momentum variables demonstrate that better price performance in the past is expected 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 in Table 6, we report the estimates of illiquidity and characteristic rewards ($\widehat{\phi}$ and $\widehat{c}_m$'s) for the excess return (*FF3-adj EXSRET2*), which are now risk-adjusted using rolling estimates of betas as described in Section 2. First, the impact of $[ILLIQ\_1]^{1/2}$ on risk-adjusted returns is virtually the same as the result with *FF3-adj EXSRET1*: positive and statistically significant at the 1% level. *SIZE* continues to have an impact on excess returns, with their statistical significance being similar. Using *FF3-adj EXSRET2*, the momentum variables are also strongly positively related to returns.

To gauge the impact of illiquidity on the stock return in the *EXSRET* specification, we find that an increase in illiquidity ($[ILLIQ\_1]^{1/2}$) by one standard deviation results in higher monthly (excess) returns of 0.23%. The magnitude of these additionally required monthly returns is economically significant, given that Chordia, Huh, and Subrahmanyam (2005) document that the average monthly (raw) return is 1.19% for 1647.2 NYSE/AMEX stocks over the past 39.5 years.

Next, we investigate in Panel B of Table 6 how the effects of illiquidity and other firm characteristics on returns change when we employ our second illiquidity measure, $[ILLIQ\_2]^{1/2}$, in the regressions. We see that the coefficients of $[ILLIQ\_2]^{1/2}$ are also statistically different from zero at the 1% level after accounting for the effects of other characteristics. As observed in Panel A, SIZE has a similar impact on returns while BTM plays no role. The momentum effects are also similar to those in Panel A.[20]

4.3 Robustness checks

4.3.1 Using quote midpoint returns. A recent study by Bessembinder and Kalcheva (2006) argues that empirical pricing tests using observed returns calculated using the reported closing prices might induce microstructure biases because of the bid-ask bounce, suggesting that asset-pricing tests with quote midpoint returns can reduce this problem. To address this issue, we obtain midpoint returns. Monthly quote midpoint returns for our sample of NYSE/AMEX stocks are calculated based on the first (open) quote midpoint and the last (close) quote midpoint (open-to-close midpoints) within each month over the 156 months from 1990 to 2002 (thirteen years: 199001–200212).[21]

The cross-sectional regression results using midpoint returns are reported in Table 7. As shown in both panels, the coefficients of $[ILLIQ\_1]^{1/2}$ and $[ILLIQ\_2]^{1/2}$ are statistically significant at 1%–5%. The slightly lower t-values of the two illiquidity measures than those in Table 6 might result from the elimination of the bid-ask bounce effect, as well as the narrower sample and the shorter time period used to compute quote midpoint returns. While the size

[20] Spiegel and Wang (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 FF factors using the data from the past sixty months (at least twenty-four months). Including SIGMA in Equation (15), we then conduct the same type of regression analysis equivalent to Table 6. The result shows that the coefficients of SIGMA are negative and statistically significant at the 5% level. However, our two theory-based illiquidity measures remain significant at the 1% level even after controlling for SIGMA. In addition to the above test, we also assessed the impact of two more filters. First, we excluded stocks with no analyst coverage, to ensure that our results are not driven by our treatment of stocks with no I/B/E/S analyst following. Second, we imposed a five-dollar price filter along the lines of Jegadeesh and Titman (2001) and Hvidkjaer (2006). The unreported results equivalent to Table 6 show that after imposing each of the two filters, the statistical significance and magnitude of the loadings on our two transformed illiquidity measures are smaller, but the coefficients remain significant at the 5% level. This suggests that our main results are not driven by very small stocks. The relevant tables are available from the authors.

[21] To obtain the quote midpoint returns for NYSE/AMEX stocks, we process the intraday quotes and trades from the Institute for the Study of Securities Markets (ISSM) and the NYSE trades and automated quotations (TAQ) databases using the Lee and Ready (1991) algorithm. For details, see Chordia, Roll, and Subrahmanyam (2002).

Table 7
Results of monthly cross-sectional regressions using quote midpoint returns: with [*ILLIQ_1*]$^{1/2}$ and [*ILLIQ_2*]$^{1/2}$ for NYSE/AMEX stocks

Panel A: Using quote midpoint returns, with square root of ILLIQ_1				Panel B: Using quote midpoint returns, with square root of ILLIQ_2			
Expla. variables	*EXSRET*	*FF3-adj EXSRET1*	*FF3-adj EXSRET2*	Expla. Variables	*EXSRET*	*FF3-adj EXSRET1*	*FF3-adj EXSRET2*
Intercept	2.851**	2.123**	2.228**	Intercept	2.975**	2.308**	2.347**
	5.75	*8.88*	*8.16*		*6.08*	*9.42*	*8.62*
[*ILLIQ_1*]$^{1/2,a}$	0.335*	0.384**	0.537**	[*ILLIQ_2*]$^{1/2,a}$	0.196**	0.150*	0.368**
	2.47	*2.75*	*2.69*		*3.04*	*2.46*	*3.37*
SIZE^{a}	−0.414**	−0.371**	−0.406**	*SIZE*^{a}	−0.428**	−0.392**	−0.419**
	−7.99	*−13.20*	*−12.47*		*−8.36*	*−13.00*	*−12.10*
BTM	−0.002	−0.057	−0.139	*BTM*	−0.033	−0.085	−0.159*
	−0.02	*−0.79*	*−1.85*		*−0.33*	*−1.18*	*−2.14*
MOM1	0.144	0.495	0.502	*MOM1*	0.232	0.574	0.627
	0.28	*1.23*	*1.03*		*0.44*	1.39	*1.26*
MOM2	1.133*	1.192**	1.046*	*MOM2*	1.122*	1.169**	1.120*
	2.39	*3.16*	*2.31*		*2.30*	*3.04*	*2.45*
MOM3	1.825**	1.076**	0.726	*MOM3*	1.846**	1.108**	0.753
	4.06	*2.89*	*1.51*		*4.05*	*2.88*	*1.53*
MOM4	1.005**	0.946**	0.577	*MOM4*	1.028**	0.943**	0.659
	2.64	*2.94*	*1.38*		*2.67*	*2.87*	*1.58*
Avg. R-sqr.	0.051	0.030	0.035	Avg. R-sqr.	0.052	0.031	0.037
Avg. obs.	1190.2	1187.4	1106.8	Avg. obs.	1127.2	1125.5	1074.6

This table reports the monthly Fama–MacBeth (1973)-type cross-sectional regressions using quote midpoint returns with [*ILLIQ_1*]$^{1/2}$ (in Panel A) and [*ILLIQ_2*]$^{1/2}$ (in Panel B) for NYSE/AMEX stocks over the 156 months (thirteen years: 199001–200212). Monthly quote midpoint returns for NYSE/AMEX stocks are calculated based on the first (open) quote midpoint and the last (close) quote midpoint (open-to-close midpoints) within each month from 1990 to 2002. The dependent variables (*EXSRET*, *FF3-adj EXSRET1*, and *FF3-adj EXSRET2*) are all one-month leading. 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; *FF3-adj EXSRET1*: the risk-adjusted excess return using the Fama–French (FF) three factors, i.e., the constant term plus the residual from the time-series regression of the excess return on the FF3 factors using the entire sample range of the data; *FF3-adj EXSRET2*: the risk-adjusted excess return using the FF3 factors with factor loadings being estimated from the five-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 sixty months in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \epsilon$, 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 FF3 factors; *ILLIQ_1* and *ILLIQ_2*: the two illiquidity measures defined in Table 1; *SIZE*: natural logarithm of *MV*, which is defined as the month-end stock price times the number of shares outstanding (in million dollars); *BTM*: natural logarithm of *BM_Trim*, which is the Winsorized book-to-market ratio, where book-to-market ratios greater than the 99.5 percentile value or less than the 0.5 percentile value in a month are set equal to the 99.5 and 0.5 percentile values, respectively; *MOM1*: compounded holding-period midpoint return of a stock over the most recent three months (from month $t-1$ to month $t-3$); *MOM2*: compounded holding-period midpoint return over the next recent three months (from month $t-4$ to month $t-6$); *MOM3*: compounded holding-period midpoint return over the three months from month $t-7$ to month $t-9$; *MOM4*: compounded holding-period midpoint return over the three months from month $t-9$ to month $t-12$. To remove nonstationarity, the Gallant, Rossi, and Tauchen (GRT) (1993) procedure has been applied to [*ILLIQ_1*]$^{1/2}$, [*ILLIQ_2*]$^{1/2}$, and *SIZE* 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. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

effect continues to be strong, the momentum effects are weakened with quote midpoint returns. We observe no book-to-market effect again.

We also perform an additional robustness check using midpoint returns for NASDAQ stocks. For these stocks, we compute returns based on the midpoints of the monthly closing quotes (close-to-close midpoints) using the NASDAQ National Market System (NMS) data in the CRSP file over the 240 months (twenty-one years) from January 1983 to December 2002. To conserve space, we do not report the corresponding table for NASDAQ stocks.[22] But the results show that the coefficients of our (transformed) illiquidity measures range from 0.124 to 0.386 and their t-values from 2.82 to 4.94. This demonstrates that the coefficients of our two measures are statistically significant at 1% for NASDAQ stocks. We also examine whether our theory-based illiquidity measures have incremental impact on returns after controlling for a direct empirical measure of illiquidity, i.e, bid-ask spreads. Given that bid and ask quotes for NASDAQ stocks are available for CRSP, we compute the quoted spread (termed *QSPRD*) as the difference between closing ask and bid quotes. Including *QSPRD* in Equation (15) for NASDAQ stocks, we then estimate the analog of Table 7 (with the dependent variable being quote midpoint returns). The results show that the coefficient of *QSPRD* is statistically significant but negative, while our two theory-based illiquidity measures remain positive and significant at the 1% level. Overall, the reliable significance of our theory-based illiquidity measures for NASDAQ stocks is reassuring because it provides out-of-sample confirmation that our theory-based measures are priced in the cross-section.

4.3.2 Different combinations of input variables in estimating the illiquidity measures. Apart from the test above using quote midpoint returns, we have already used three different types of excess returns (*EXSRET*, *FF3-adj EXSRET1*, and *FF3-adj EXSRET2*) to ensure the robustness of our results.[23] Now we consider the effects of our choices of input variables used to estimate the two illiquidity measures. As pointed out in Section 1, the two important input variables in estimating $[ILLIQ\_1]^{1/2}$ and $[ILLIQ\_2]^{1/2}$ are, among others, the standard deviation of daily returns in month $t-1$ and the average of daily dollar volume in month $t-1$. For further robustness, we obtain standard deviations of returns computed each month with daily returns in month $t-2$, daily returns in the past thirty-six months, and monthly returns in the past sixty months. Moreover, as a proxy for $std(R)$, we also use an idiosyncratic risk measure against the FF three factors using data from the past sixty months, as in Spiegel and Wang (2005). For the average volume as a proxy for

[22] Month-end bid and ask quotes for the NASDAQ National Market System (NMS) stocks are obtained using the relevant CRSP files. The full results corresponding to Table 7 for NASDAQ stocks are available from the authors.

[23] As mentioned earlier, we additionally obtained a "purged" estimator of Black, Jensen, and Scholes (1972) for each of the explanatory variables in the regressions of *FF3-adj EXSRET1* and *FF3-adj EXSRET2*. 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.

$std(z)$, we try many candidates, including average daily share volume, average daily dollar volume, and daily turnover in month $t-1$, $t-2$ and in the past thirty-six months. We estimate the two illiquidity measures using a number of combinations of these as inputs. Cross-sectional regressions using illiquidity measures estimated with different combinations of input variables do not significantly change our results, especially the effects of illiquidity, firm size, and book-to-market equity.

Thus far, we have demonstrated that the two theory-based illiquidity measures continue to be priced in the cross-section of stock returns, regardless of using quote midpoint returns, as well as different input variables in estimating the illiquidity measures. However, three more questions still remain to be answered: (i) do the theory-based measures perform better than the other commonly used (il)liquidity measures in the finance literature? (ii) Do the two measures continue to be priced after accounting for the effects of other competing (il)liquidity measures? (iii) Given the high skewness in our square-root transformed measures, is the illiquidity pricing observed thus far driven primarily by a few outlier stocks? To further test the robustness of our findings, we compare the results of the theory-based measures to those of alternative illiquidity measures. We examine the issue that may arise from the high skewness of our illiquidity measures in the next section.

5. A Horse Race with Alternative Measures

5.1 Selection of alternative measures and their relations to the theory-based illiquidity measures

There are a number of (il)liquidity measures that have been used in the asset-pricing or microstructure literature. Some measures have been obtained or estimated from the TAQ database, while others come from CRSP data. As Hasbrouck (2005) admits, however, estimating the measures using high-frequency trade and quote data, such as the TAQ database, limits the availability to the relatively small and recent data samples. Merton (1980) 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, and hence the importance, of liquidity measures that can be constructed from data of daily or lower frequency. Of course, our two theory-based measures can be constructed from the CRSP daily file and the lower frequency I/B/E/S database. 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, and thus over the same horizons as our measures.

First, we consider Amihud's (2002) illiquidity measure, which is defined as $|r|/DDVOL$, where r is the daily stock return and $DDVOL$ is the daily dollar volume (in thousand dollars). 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 (λ), which is the basis of our two theory-based illiquidity measures. However, the Amihud measure is distinct from Kyle's λ in the sense that *Amihud* captures the absolute return impact of *unsigned* volume, while λ is the price impact of *signed* volume (order flows). From an operational standpoint, our closed-form expressions for lambdas include the impact of analyst following, over and above the volatility and volume measures, since it is an economic link between illiquidity and information flows. Given the fact that the Amihud measure has been used widely in the 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 the appropriate CRSP-based proxies for a TAQ-based effective cost. We thus consider this measure (notated as *Roll_Gibbs*) in our study.[24]

From a practitioner's perspective, information about dollar volume may be more important than that about share volume or turnover. Moreover, Spiegel and Wang (2005) argue that the only liquidity measure that explains cross-sectional returns above and beyond the other variables is the dollar volume. To look into this issue further, we include the monthly average of daily dollar volume (in thousand dollars) within each month for each stock (notated as *DVOL*) as one of the liquidity proxies in our analysis.

Lastly, share turnover has been used as an (il)liquidity proxy by many researchers, including Brennan, Chordia, and Subrahmanyam (1998). For this reason, we consider the monthly average of daily share turnover within each month (notated as *TURN*) as an alternative to *DVOL* described above. The three (il)liquidity measures (*Amihud*, *DVOL*, and *TURN*) are estimated each month and the *Roll_Gibbs* measure is available at an annual frequency.

Table 8 presents the correlation coefficients between the (il)liquidity measures. Panel A shows that both *ILLIQ_1* and *ILLIQ_2* are most highly correlated with *Amihud*, followed by *Roll_Gibbs*. Given that high dollar volume and turnover are positively associated with liquidity, both *DVOL* and *TURN* are negatively correlated with the two theory-based illiquidity measures. We also find that *Amihud* is highly correlated with *Roll_Gibbs*. An interesting aspect is that the correlation coefficient between *TURN* and *DVOL* is not so high, although the two may often be considered as an alternative to each other. Panel B is the analog of Panel A, using all the transformed measures. With the square-root transforms, the (absolute) correlation coefficients increase substantially, but the patterns are qualitatively the same as in Panel A.

[24] This measure is obtained from Joel Hasbrouck's website, and is described in Hasbrouck (2006), where the measure is denoted by *c_BMA*.

Table 8
Relations of the theory-based illiquidity measures to alternative measures for NYSE/AMEX stocks

Panel A: Average correlations between (il)liquidity measures

	ILLIQ_1	*ILLIQ_2*	*Amihud*	*Roll_Gibbs*	*DVOL*	*TURN*
ILLIQ_1	1					
ILLIQ_2	0.604	1				
Amihud	0.650	0.460	1			
Roll_Gibbs	0.498	0.321	0.552	1		
DVOL	−0.053	−0.030	−0.055	−0.171	1	
TURN	−0.087	−0.047	−0.092	−0.109	0.215	1

Panel B: Average correlations between square-root values of (il)liquidity measures

	$[ILLIQ\_1]^{1/2}$	$[ILLIQ\_2]^{1/2}$	$[Amihud]^{1/2}$	$[Roll\_Gibbs]^{1/2}$	$[DVOL]^{1/2}$	*TURN*
$[ILLIQ\_1]^{1/2}$	1					
$[ILLIQ\_2]^{1/2}$	0.773	1				
$[Amihud]^{1/2}$	0.864	0.694	1			
$[Roll\_Gibbs]^{1/2}$	0.693	0.531	0.711	1		
$[DVOL]^{1/2}$	−0.321	−0.210	−0.321	−0.405	1	
$[TURN]^{1/2}$	−0.283	−0.183	−0.299	−0.169	0.411	1

Panel A shows the monthly average correlations between the key (il)liquidity measures for NYSE/AMEX stocks over the past 324 months (twenty-seven years: 197601–200212), while Panel B reports the average correlations between the square-root values of the six corresponding (il)liquidity measures. 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: *ILLIQ_1*: the first illiquidity measure defined as in Table 1; *ILLIQ_2*: the second illiquidity measure defined as in Table 1; *Amihud*: the illiquidity measure of Amihud (2002) estimated each month as the average of $|r|/DDVOL$, where r is the daily stock return and DDVOL is the daily dollar volume in thousands; *Roll_Gibbs*: the market risk-adjusted effective bid-ask spread of Roll (1984) estimated using the Gibbs sampler, which is of annual frequency obtained from the website of Joel Hasbrouck; *DVOL*: the average of daily dollar volume within each month for each stock (in thousands); *TURN*: the average of daily share turnover within each month for each stock. The average number of component stocks used in a year is 1826.0.

5.2 Cross-sectional regressions with alternative illiquidity measures

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

First, we run the regression with each of the competing (square-root transformed) measures by replacing Γ_i_{jt} in Equation (15) with one of the three measures ($[Amihud]^{1/2}$, $[Roll\_Gibbs]^{1/2}$, and either $[DVOL]^{1/2}$ or $[TURN]^{1/2}$).[25] The three alternative measures are GRT-adjusted before conducting the cross-sectional regressions. For brevity, these results are not tabulated, but they are summarized as follows. As the correlation coefficients in Table 8 suggest, the impact of $[Amihud]^{1/2}$ on returns is consistently strong. However, the impact

[25] Since $[Roll\_Gibbs]^{1/2}$ 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.

of $[Roll\_Gibbs]^{1/2}$ has the wrong sign (negative), although it is statistically significant. This is surprising, given that this measure is highly correlated with $[Amihud]^{1/2}$, as well as with the two theory-based measures. As Hasbrouck (2006) indicates, the limitation of $[Roll\_Gibbs]^{1/2}$ stems from the fact that it does not explicitly incorporate the price impact effects of trading volume or order flow, which may be endogenous with price dynamics. The negative sign on $[Roll\_Gibbs]^{1/2}$ is consistent with Eleswarapu and Reinganum (1993), who uncovered a similar result for bid-ask spread in the 1981–1990 period (Panel B of their Table 4).

As one would expect, dollar volume and firm size are highly correlated. A preliminary test shows that the coefficients of cross-sectional correlation (each month) between GRT-adjusted $[DVOL]^{1/2}$ and SIZE are often higher than 70%, while those of time-series correlation (for the whole sample period for each firm) between the two variables are often higher than 85%. This causes a multicollinearity problem when we include $[DVOL]^{1/2}$ and SIZE at the same time in the regression, in which case the coefficient of $[DVOL]^{1/2}$ is significant, but its sign is positive. When we exclude SIZE in the regression, however, $[DVOL]^{1/2}$ is negative and statistically significant at any conventional level (also see the similar patterns in Panels A and B of Table 9). On the contrary, the impact of $[TURN]^{1/2}$ on returns is negligible after controlling for other firm characteristics (whether SIZE is included or not).[26]

Now, we run a horse race between one of our two theory-based illiquidity measures and all the three competing measures together. For this purpose, we augment Equation (15) by including three more variables as in the equation

$$\tilde{R}^{*}_{jt+1} = c_{0t} + \phi_t \Gamma\_i_{jt} + \sum_{n=1}^{3} \varphi_{nt} ALT_{njt} + \sum_{m=1}^{M} c_{mt} Z_{mjt} + \tilde{e}'_{jt+1}, \text{ i=1 or 2,} \tag{16}$$

where ALT_{njt} $(n = 1, \ldots, 3)$ denotes one of the three alternative (il)liquidity measures ($[Amihud]^{1/2}$, $[Roll\_Gibbs]^{1/2}$, and either $[DVOL]^{1/2}$ or $[TURN]^{1/2}$). By including the three additional (il)liquidity measures in the regressions, as shown in Panels A–D of Table 9, the adjusted R^2 increases, compared to the relevant specifications reported in Table 6.

The results with $[DVOL]^{1/2}$ are reported in Panels A and B of Table 9. These panels address whether each of the two theory-based illiquidity measures have any incremental impact on returns after accounting for the effects of alternative measures including dollar volume, $[DVOL]^{1/2}$. Because of the multicollinearity problem mentioned above, two different specifications are reported for each of the two risk-adjusted excess returns (*FF3-adj EXSRET1*

[26] An alternative measure of illiquidity is the well-known reversal measure of Pástor and Stambaugh (PS, 2003). An unreported result shows that the impact of this measure is modest after controlling for other firm characteristics. This is consistent with Hasbrouck (2005, see his Table 5). Unlike us, however, PS use their measure to price illiquidity risk, rather than the level of illiquidity.

Table 9
A horse race with all the (il)liquidity measures together for NYSE/AMEX stocks

Panel A: With square root of ILLIQ_1 and 3 alternative measures incl. $[DVOL]^{1/2}$

	FF3-adj EXSRET1		*FF3-adj EXSRET2*	
Expla. var.	1	2	3	4
Intercept	0.935**	3.808**	1.014**	3.971**
	6.34	*18.57*	*6.73*	*19.64*
$[ILLIQ\_1]^{1/2,a}$	0.392**	0.277**	0.382**	0.267**
	7.68	*5.47*	*7.16*	*4.98*
$[Amihud]^{1/2,a}$	6.463**	4.599**	7.658**	5.666**
	6.84	*4.91*	*7.47*	*5.64*
$[Roll\_Gibbs]^{1/2,a}$	−16.654**	−23.457**	−17.092**	−24.082**
	−10.96	*−15.58*	*−10.84*	*−15.51*
$[DVOL]^{1/2,a}$	−0.023**	0.031**	−0.025**	0.030**
	−8.52	*8.17*	*−9.25*	*7.43*
$SIZE^a$		−0.482**		−0.496**
		−18.68		*−18.06*
BTM	0.158**	0.002	0.107*	−0.055
	3.38	*0.04*	*2.30*	*−1.20*
*MOM*1	0.389	0.572*	0.444	0.640*
	1.42	*2.10*	*1.43*	*2.08*
MOM2	0.730**	0.908**	0.759**	0.944**
	3.06	*3.83*	*2.83*	*3.53*
MOM3	0.733**	0.900**	0.460	0.633*
	2.96	*3.65*	*1.59*	*2.21*
MOM4	0.644**	0.739**	0.409	0.514*
	3.24	*3.69*	*1.62*	*2.02*
Avg. R-sqr.	0.034	0.037	0.036	0.039
Avg. obs.	1715.3	1715.3	1676.7	1676.7

Panel B: With square root of ILLIQ_2 and 3 alternative measures incl. $[DVOL]^{1/2}$

	FF3-adj EXSRET1		*FF3-adj EXSRET2*	
Expla. var.	5	6	7	8
Intercept	0.890**	3.818**	0.975**	4.027**
	6.08	*18.48*	*6.46*	*19.87*
$[ILLIQ\_2]^{1/2,a}$	0.103**	0.094**	0.108**	0.099**
	4.16	*3.78*	*4.25*	*3.89*
$[Amihud]^{1/2,a}$	9.394**	5.924**	10.292**	6.646**
	10.75	*6.89*	*11.31*	*7.52*
$[Roll\_Gibbs]^{1/2,a}$	−14.968**	−22.319**	−15.560**	−23.208**
	−9.68	*−14.70*	*−9.79*	*−14.92*
$[DVOL]^{1/2,a}$	−0.025**	0.029**	−0.026**	0.030**
	−9.19	*7.63*	*−9.59*	*7.35*
$SIZE^a$		−0.486**		−0.507**
		−18.64		*−18.59*
BTM	0.148**	−0.021	0.100*	−0.076
	3.16	*−0.45*	*2.15*	*−1.64*
MOM1	0.360	0.556*	0.449	0.663*
	1.31	*2.04*	*1.47*	*2.19*
MOM2	0.692**	0.885**	0.771**	0.972**
	2.88	*3.71*	*2.88*	*3.64*
MOM3	0.704**	0.874**	0.464	0.641*
	2.82	*3.51*	*1.59*	*2.22*
MOM4	0.597**	0.691**	0.343	0.447
	2.91	*3.35*	*1.34*	*1.73*
Avg. R-sqr.	0.036	0.040	0.038	0.041
Avg. obs.	1607.6	1607.6	1596.5	1596.5

Table 9
Continued

Panel C: With square root of ILLIQ_1 and 3 alternative measures incl. $[TURN]^{1/2}$

	FF3-adj EXSRET1		*FF3-adj EXSRET2*	
Expla. var.	9	10	11	12
Intercept	0.808**	3.298**	0.874**	3.471**
	4.86	*15.31*	*5.30*	*16.52*
$[ILLIQ\_1]^{1/2,a}$	0.405**	0.306**	0.395**	0.294**
	7.91	*6.03*	*7.38*	*5.48*
$[Amihud]^{1/2,a}$	6.515**	4.952**	7.685**	6.012**
	6.75	*5.19*	*7.41*	*5.90*
$[Roll\_Gibbs]^{1/2,a}$	−15.705**	−22.874**	−15.969**	−23.462**
	−10.63	*−15.11*	*−10.40*	*−15.01*
$[TURN]^{1/2,a}$	−1.906	1.284	−2.072	1.248
	−1.14	*0.76*	*−1.19*	*0.71*
$SIZE^a$		−0.376**		−0.393**
		−19.92		*−20.07*
BTM	0.205**	0.006	0.160**	−0.052
	4.48	*0.12*	*3.50*	*−1.14*
MOM1	0.445	0.645*	0.501	0.716*
	1.64	*2.39*	*1.64*	*2.34*
MOM2	0.724**	0.929**	0.755**	0.969**
	3.06	*3.92*	*2.83*	*3.62*
MOM3	0.681**	0.872**	0.413	0.608*
	2.77	*3.55*	*1.44*	*2.11*
MOM4	0.614**	0.728**	0.379	0.507*
	3.14	*3.68*	*1.51*	*2.00*
Avg. R-sqr.	0.035	0.039	0.038	0.041
Avg. obs.	1715.3	1715.3	1676.7	1676.7

Panel D: With square root of ILLIQ_2 and 3 alternative measures incl. $[TURN]^{1/2}$

	FF3-adj EXSRET1		*FF3-adj EXSRET2*	
Expla. var.	13	14	15	16
Intercept	0.785**	3.349**	0.848**	3.541**
	4.81	*15.57*	*5.21*	*16.86*
$[ILLIQ\_2]^{1/2,a}$	0.103**	0.096**	0.108**	0.101**
	4.14	*3.86*	*4.24*	*3.99*
$[Amihud]^{1/2,a}$	9.573**	6.546**	10.467**	7.273**
	10.59	*7.38*	*11.17*	*7.97*
$[Roll\_Gibbs]^{1/2,a}$	−13.851**	−21.630**	−14.316**	−22.475**
	−9.17	*−14.12*	*−9.25*	*−14.30*
$[TURN]^{1/2,a}$	−2.868	0.558	−2.729	0.824
	−1.69	*0.33*	*−1.58*	*0.47*
$SIZE^a$		−0.385**		−0.404
		−20.16		*−20.64*
BTM	0.199**	−0.018	0.156**	−0.074
	4.34	*−0.41*	*3.40*	*−1.61*
MOM1	0.405	0.620*	0.494	0.727*
	1.48	*2.29*	*1.64*	*2.42*
MOM2	0.683**	0.903**	0.764**	0.994**
	2.86	*3.78*	*2.86*	*3.72*
MOM3	0.647**	0.845**	0.415	0.616*
	2.62	*3.40*	*1.43*	*2.12*
MOM4	0.563**	0.679**	0.310	0.439
	2.79	*3.32*	*1.22*	*1.70*
Avg. R-sqr.	0.037	0.041	0.039	0.042
Avg. obs.	1607.6	1607.6	1596.5	1596.5

(*continued overleaf*)

Table 9
Continued

This table runs a horse race in the monthly Fama–MacBeth (1973)-type cross-sectional regressions using one of our illiquidity measures together with the three alternative (il)liquidity measures for NYSE/AMEX stocks over the past 324 months (twenty-seven years: 197601–200212). Panel A contains the results for $[ILLIQ\_1]^{1/2}$ with the three alternative measures including the square root of dollar volume, $[DVOL]^{1/2}$, while Panel B does the same for $[ILLIQ\_2]^{1/2}$. Panels C and D are the analogs to Panels A and B, but with the square root of share turnover, $[TURN]^{1/2}$, instead of $[DVOL]^{1/2}$. The dependent variables (*FF3-adj EXSRET1* and *FF3-adj EXSRET2*) are all one-month leading. The definitions of the variables are as follows: *FF3-adj EXSRET1*: the risk-adjusted excess return using the Fama–French (FF) three factors, i.e., the constant term plus the residual from the time-series regression of the excess return on the FF3 factors using the entire sample range of the data; *FF3-adj EXSRET2*: the risk-adjusted excess return using the FF3 factors with factor loadings being estimated from the five-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 sixty months in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \epsilon$, 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 FF3 factors; $[ILLIQ\_1]^{1/2}$: square root of ILLIQ_1, which is defined as $\frac{N^{0.5} std(R)}{(N+1) std(z)}$, where N is the number of informed traders, *std(R)* is the standard deviation of returns, and *std(z)* is the standard deviation of noise trades (the original input variables are proxied by the variables as shown in Table 1); $[ILLIQ\_2]^{1/2}$: square root of ILLIQ_2, which is defined as $P^{-1} \frac{\nu_\delta}{(N+1)\nu_\delta + 2\nu_\epsilon} \sqrt{\frac{N(\nu_\delta + \nu_\epsilon)}{\nu_z}}$, where P is the asset price, N is the number of informed traders, ν_δ is the variance of payoff innovations, ν_ϵ is the variance of signal innovations, and ν_z is the variance of noise trades (the original input variables are proxied by the variables as shown in Table 1); $[Amihud]^{1/2}$: square root of Amihud, which is the illiquidity measure of Amihud (2002) estimated each month as the average of $|r|/DDVOL$, where r is the daily stock return and *DDVOL* is the daily dollar volume in thousands; $[Roll\_Gibbs]^{1/2}$: square root of Roll_Gibbs, which is the market risk-adjusted effective bid-ask spread of Roll (1984) estimated using the Gibbs sampler (obtained from the website of Joel Hasbrouck); $[DVOL]^{1/2}$: square root of DVOL, which is defined as the average of daily dollar volume within each month for each stock (in thousands); $[TURN]^{1/2}$: square root of *TURN*, which is the average of daily share turnover within each month for each stock; *SIZE*: natural logarithm of *MV*, which is defined as the month-end stock price times the number of shares outstanding (in million dollars); *BTM*: natural logarithm of *BM_Trim*, which is the Winsorized book-to-market ratio, where book-to-market ratios greater than the 99.5 percentile value or less than the 0.5 percentile value in a month are set equal to the 99.5 and 0.5 percentile values, respectively; *MOM1*: compounded holding-period return of a stock over the most recent three months (from month $t-1$ to month $t-3$); *MOM2*: compounded holding-period return over the next recent three months (from month $t-4$ to month $t-6$); *MOM3*: compounded holding-period return over the three months from month $t-7$ to month $t-9$; *MOM4*: compounded holding-period return over the three months from month $t-9$ to month $t-12$. For monthly regressions, we keep the annual values of *Roll_Gibbs* constant over the twelve months within each year. To remove nonstationarity, the Gallant, Rossi, and Tauchen (GRT, 1993) procedure has been applied to $[ILLIQ\_1]^{1/2}$, $[ILLIQ\_2]^{1/2}$, $[Amihud]^{1/2}$, $[Roll\_Gibbs]^{1/2}$, $[DVOL]^{1/2}$, $[TURN]^{1/2}$, and *SIZE* 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. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

and *FF3-adj EXSRET2*) in each panel (i.e., specifications 1, 3, 5, and 7 exclude *SIZE*, but specifications 2, 4, 6, and 8 include it). The results from the regression specifications with the risk-unadjusted returns (*EXSRET*) are not reported.

To briefly summarize the results presented in Panels A and B of Table 9, we find that $[DVOL]^{1/2}$ is strongly negatively related to returns in the regression specifications 1, 3, 5, and 7 that exclude *SIZE*. In specifications 2, 4, 6, and 8, however, the sign of the loadings on $[DVOL]^{1/2}$ reverses because of multicollinearity with *SIZE*. In both panels, the sign of the loadings on the other two alternative measures ($[Amihud]^{1/2}$ and $[Roll\_Gibbs]^{1/2}$) and their statistical significance are essentially the same as those described above. In any case, however, the coefficients of our two illiquidity measures are statistically significant at any conventional level even after controlling for dollar volume, together with other illiquidity measures, as well as firm characteristics. The strong role of dollar volume is consistent with Spiegel and Wang (2005). Other interesting features are: without *SIZE* the book-to-market effect is more evident, but the momentum effects turn weaker.

Next, we turn to the regressions with share turnover, $[TURN]^{1/2}$, as an alternative to dollar volume, $[DVOL]^{1/2}$. As Panels C and D of Table 9 show, the patterns with turnover are qualitatively similar to those in Panels A and B. The Amihud measure is consistently positively related to returns, and the Roll measure continues to have the wrong sign. The only difference is that the coefficients of $[TURN]^{1/2}$ are not statistically significant. In any case, our two illiquidity measures are strongly positively related to returns.

We now consider the issue that even our transformed measures are a bit left-skewed (see Table 2). This raises the concern that illiquidity pricing may be driven by a few small stocks with large values of illiquidity. A simple and intuitive way to address this issue is to exclude such stocks before conducting the regressions. For this purpose, each month the component stocks are split, in turn, into size and illiquidity quintiles (by each of our two measures). Then the stocks that belong to the largest illiquidity quintile or the smallest size quintile are excluded before running a horse race. By applying the above filter, the number of monthly component stocks decreases on average by 423.6–434.5 across the different specifications. The results are shown in Table 10. As can be seen from the regressions, the statistical significance of our measures is largely unaltered, while the size of their coefficients increases relative to that in Table 9. Intriguingly, the role of the Amihud measure is limited in Table 10, suggesting that this measure works best for the most illiquid stocks. Overall, Table 10 confirms that the role of our illiquidity measures in the cross-section is not driven by a few very small, illiquid stocks.

To sum up, our empirical tests provide evidence that theory-based illiquidity is a priced attribute in the cross-section of expected returns, even after controlling for other illiquidity measures prevalent in the literature.

Table 10

A horse race after excluding illiquidity quintile 5 (highest illiquidity group) and size quintile 1 (smallest size group) for NYSE/AMEX stocks

Excluding illiquidity quintile 5 (highest illiquidity group) and size quintile 1 (smallest size group)

Panel A: With square root of ILLIQ_1 and 3 alternative measures incl. $[DVOL]^{1/2}$

	FF3-adj EXSRET1		*FF3-adj EXSRET2*	
Expla. var.	1	2	3	4
Intercept	0.858**	2.880**	0.904**	2.922**
	6.14	*14.13*	*6.23*	*14.10*
$[ILLIQ\_1]^{1/2,a}$	1.094**	0.759**	1.046**	0.718**
	8.32	*5.72*	*7.70*	*5.24*
$[Amihud]^{1/2,a}$	0.805	−2.884	2.799	−0.950
	0.40	*−1.43*	*1.31*	*−0.44*
$[Roll\_Gibbs]^{1/2,a}$	−19.492**	−22.954**	−19.747**	−23.210**
	−13.56	*−16.08*	*−13.16*	*−15.60*
$[DVOL]^{1/2,a}$	−0.018**	0.010**	−0.019**	0.008*
	−6.93	*2.78*	*−7.49*	*2.13*
$SIZE^a$		−0.312**		−0.312**
		−12.95		*−12.04*
BTM	0.051	−0.014	−0.023	−0.090
	1.16	*−0.32*	*−0.50*	*−1.93*
MOM1	0.983**	1.067**	1.043**	1.137**
	3.42	*3.72*	*3.23*	*3.53*
MOM2	1.239**	1.308**	1.198**	1.268**
	4.62	*4.88*	*4.00*	*4.24*
MOM3	1.487**	1.560**	1.222**	1.303**
	5.88	*6.16*	*4.25*	*4.55*
MOM4	0.981**	1.001**	0.712*	0.744**
	4.37	*4.45*	*2.56*	*2.66*
Avg. R-sqr.	0.034	0.036	0.037	0.039
Avg. obs.	1280.8	1280.8	1251.8	1251.8

Panel B: With square root of ILLIQ_2 and 3 alternative measures incl. $[DVOL]^{1/2}$

	FF3-adj EXSRET1		*FF3-adj EXSRET2*	
Expla. var.	5	6	7	8
Intercept	0.850**	3.019**	0.913**	3.112**
	6.07	*15.32*	*6.31*	*15.65*
$[ILLIQ\_2]^{1/2,a}$	0.730**	0.665**	0.841**	0.774**
	8.47	*7.73*	*8.71*	*8.09*
$[Amihud]^{1/2,a}$	0.796	−5.358**	1.626	−4.650*
	0.46	*−2.92*	*0.87*	*−2.35*
$[Roll\_Gibbs]^{1/2,a}$	−16.923**	−21.057**	−17.670**	−21.834**
	−11.54	*−14.39*	*−11.91*	*−14.72*
$[DVOL]^{1/2,a}$	−0.022**	0.008*	−0.023**	0.007*
	−8.80	*2.30*	*−8.94*	*2.06*
$SIZE^a$		−0.336**		−0.341**
		−14.63		*−14.05*
BTM	0.040	−0.038	−0.033	−0.111*
	0.88	*−0.84*	*−0.68*	*−2.32*
MOM1	0.855**	0.952**	0.950**	1.056**
	2.97	*3.30*	*2.99*	*3.32*
MOM2	1.157**	1.241**	1.151**	1.235**
	4.28	*4.59*	*3.91*	*4.19*
MOM3	1.370**	1.460**	1.120**	1.212**
	5.38	*5.73*	*3.88*	*4.21*
MOM4	1.071**	1.108**	0.770**	0.822**
	4.63	*4.77*	*2.72*	*2.89*
Avg. R-sqr.	0.036	0.038	0.038	0.040
Avg. obs.	1180.8	1180.8	1172.9	1172.9

This table runs a horse race after excluding all stocks belong to illiquidity quintile 5 (a stock group with highest values in the two theory-based illiquidity measures) or size quintile 1 (a stock group with smallest size) in order to eliminate the possibility that illiquidity pricing is driven purely by a few outlier stocks. Each month the component stocks are split into five portfolios (illiquidity quintiles 1–5) after being sorted in ascending order by each of the two theory-based illiquidity measures ($[ILLIQ\_1]^{1/2}$ and $[ILLIQ\_2]^{1/2}$), and illiquidity group numbers 1–5 are assigned to the component stocks for each of illiquidity quintiles 1–5, respectively. Similarly, each month the component stocks are split into five portfolios (size quintiles 1–5) after being sorted in ascending order by market capitalization (*MV* in million dollars), and size group numbers 1–5 are assigned to the component stocks for each of size quintiles 1–5, respectively. Then the stocks that have illiquidity group number 5 (the highest illiquidity group) or size group number 1 (the smallest size group) are all excluded before conducting the monthly Fama–MacBeth (1973)-type cross-sectional regressions, which use one of our two illiquidity measures together with the three alternative (il)liquidity measures for NYSE/AMEX stocks over the past 324 months (twenty-seven years: 197601–200212). Panel A contains the results for $[ILLIQ\_1]^{1/2}$ with the three alternative measures including the square root of dollar volume, $[DVOL]^{1/2}$, while Panel B does the same for $[ILLIQ\_2]^{1/2}$. The dependent variables (*FF3-adj EXSRET1* and *FF3-adj EXSRET2*) are all one-month leading. The definitions of the variables are as follows: *FF3-adj EXSRET1*: the risk-adjusted excess return using the Fama–French (FF) three factors, i.e., the constant term plus the residual from the time-series regression of the excess return on the FF3 factors using the entire sample range of the data; *FF3-adj EXSRET2*: the risk-adjusted excess return using the FF3 factors with factor loadings being estimated from the five-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 (α, $\beta_1 \beta_2$, β_3) are first estimated for each month using the time-series data of the past sixty months in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \epsilon$, 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 FF3 factors; $[ILLIQ\_2]^{1/2}$: square root of *ILLIQ_1*, which is defined as $\frac{N^{0.5} std(R)}{(N+1) std(z)}$, where *N* is the number of informed traders, *std(R)* is the standard deviation of returns, and *std(z)* is the standard deviation of noise trades (the original input variables are proxied by the variables as shown in Table 1); $[ILLIQ\_2]^{1/2}$: square root of *ILLIQ_2*, which is defined as $P^{-1} \frac{\nu_\delta}{(N+1)\nu_\delta + 2\nu_\epsilon} \sqrt{\frac{N(\nu_\delta + \nu_\epsilon)}{\nu_z}}$, where *P* is the asset price, *N* is the number of informed traders, ν_δ is the variance of payoff innovations, ν_ϵ is the variance of signal innovations, and ν_z is the variance of noise trades (the original input variables are proxied by the variables as shown in Table 1); $[Amihud]^{1/2}$: square root of *Amihud*, which is the illiquidity measure of Amihud (2002) estimated each month as the average of $|r|/DDVOL$, where *r* is the daily stock return and *DDVOL* is the daily dollar volume in thousands; $[Roll\_Gibbs]^{1/2}$: square root of *Roll_Gibbs*, which is the market risk-adjusted effective bid-ask spread of Roll (1984) estimated using the Gibbs sampler (obtained from the website of Joel Hasbrouck); $[DVOL]^{1/2}$: square root of *DVOL*, which is defined as the average of daily dollar volume within each month for each stock (in thousands); *SIZE*: natural logarithm of *MV*, which is defined as the month-end stock price times the number of shares outstanding (in million dollars); *BTM*: natural logarithm of *BM_Trim*, which is the Winsorized book-to-market ratio, where book-to-market ratios greater than the 99.5 percentile value or less than the 0.5 percentile value in a month are set equal to the 99.5 and 0.5 percentile values, respectively; *MOM1*: compounded holding-period return of a stock over the most recent three months (from month $t-1$ to month $t-3$); *MOM2*: compounded holding-period return over the next recent three months (from month $t-4$ to month $t-6$); *MOM3*: compounded holding-period return over the three months from month $t-7$ to month $t-9$; *MOM4*: compounded holding-period return over the three months from month $t-9$ to month $t-12$. For monthly regressions, we keep the annual values of *Roll_Gibbs* constant over the twelve months within each year. To remove nonstationarity, the Gallant, Rossi, and Tauchen (GRT, 1993) procedure has been applied to $[ILLIQ\_1]^{1/2}$, $[ILLIQ\_2]^{1/2}$, $[Amihud]^{1/2}$, $[Roll\_Gibbs]^{1/2}$, $[DVOL]^{1/2}$, and *SIZE* 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. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

6. Conclusion

Several empirical proxies for illiquidity have been used in the literature, which links trading frictions to asset prices. These measures, however, have been subject to controversy because they have achieved mixed results when answering the question of whether illiquidity is related to asset returns. This raises the question of whether the informal empirical reasoning that justifies the various measures is the cause of conflicting conclusions about the illiquidity-return relation.

We use an alternative approach to measuring illiquidity. Specifically, we explicitly model the functional relation between illiquidity and its primitive drivers, and thus provide stronger economic underpinnings for the estimation of illiquidity relative to those in the extant literature. We estimate Kyle lambdas, using analytic formulae that are derived from an equilibrium framework. We use plausible empirical proxies for inputs to the theoretical expressions along the lines of Brennan and Subrahmanyam (1995). Our lambdas are estimated for a comprehensive sample of stocks spanning twenty-seven years of data. Asset-pricing regressions lend support to the notion that our illiquidity measures are priced in the cross-section of returns, even after accounting for the effects of other, empirically motivated, illiquidity proxies, and this phenomenon is not driven by a few outlier stocks.

Appendix: Derivation of the Two Theory-Based Illiquidity Measures

In this section, we derive the illiquidity measure (λ) used in our analysis, assuming that there are many informed traders in Kyle's (1985) setting. We begin our analysis by stating a few standard assumptions that are made in much of the literature on Kyle (1985)-type frameworks. Consider an asset that pays off $\tilde{W} = \overline{W} + \tilde{\delta}$, where $\tilde{W}$ is the liquidation value of the asset (or the common value that all traders assign to it), $\overline{W}$ is the expected value of the asset, and $\tilde{\delta}$ is the innovation in the asset payoff that is normally distributed (with its mean being zero), i.e., $\tilde{\delta} \sim N(0, v_\delta)$. There are N informed traders who observe a signal that is informative about $\tilde{\delta}$. For now, we assume that informed trader i observes a signal with an error, $\tilde{\delta} + \tilde{\varepsilon_i}$ $(i = 1, 2, 3, \ldots, N)$, where $\tilde{\varepsilon_i}'$s are *iid* and normally distributed, i.e., $\tilde{\varepsilon_i}$ $\sim$*iid* $N(0, v_\varepsilon)$. Informed traders maximize expected profits. There are also uninformed traders who trade randomly, and their total trades, z, are normally distributed, i.e., $z \sim N(0, v_z)$. It is assumed that δ, ε, and z are all independent. Risk-neutral market makers set the prices of assets equal to the expected values of the liquidation values, conditional on information about the quantities traded by other participants. They are competitive and efficient, earning zero expected profits and ensuring that markets are clear.

At each auction, the trading of an asset occurs in two steps. In the first step, the informed and uninformed traders submit orders simultaneously to a market maker. In the second step, the market maker quotes a price contingent on the *combined* trades (order flows) of both types of traders. The market maker does not observe the *individual* quantities traded by the informed or the uninformed. The individual does not have any other information than the combined total trades by the two types of traders. Therefore, price fluctuations of an asset are purely a result of order flow innovations.

Suppose that informed trader i conjectures that other informed traders use trading strategies of a form $\gamma(\tilde{\delta} + \tilde{\varepsilon}_j)$, i.e., a trade of informed trader j is given by

$$x_j = \gamma(\tilde{\delta} + \tilde{\varepsilon}_j), \tag{A1}$$

and also that for informed trader i is

$$x_i = \gamma(\tilde{\delta} + \tilde{\varepsilon}_i). \tag{A2}$$

From the above equations, the combined total trades (order flows), ω, are expressed as a sum of informed and uninformed trades, i.e.,

$$\omega = \{x_i + (N-1)x_j\} + z$$

$$= \left\{ x_i + (N-1)\gamma\tilde{\delta} + \gamma \sum_{j \neq i} \tilde{\varepsilon}_j \right\} + z \tag{A3}$$

$$= N\gamma\tilde{\delta} + \gamma \sum_i \tilde{\varepsilon_i} + z. \tag{A4}$$

The asset price, P, is set by the market maker after the individual observes ω so that[27]

$$P = E\left[\overline{W} + \tilde{\delta} \mid \omega = N\gamma\tilde{\delta} + \gamma \sum_i \tilde{\varepsilon_i} + z \right] \tag{A5}$$

$$= \overline{W} + \frac{Cov\left[\tilde{\delta}, N\gamma\tilde{\delta} + \gamma \sum_i \tilde{\varepsilon_i} + z\right]}{Var\left[N\gamma\tilde{\delta} + \gamma \sum_i \tilde{\varepsilon_i} + z\right]} \left(N\gamma\tilde{\delta} + \gamma \sum_i \tilde{\varepsilon_i} + z \right). \tag{A6}$$

In addition, Kyle (1985) suggests that P should also be a linear function of order flows in a form,

$$P = \overline{W} + \lambda\omega \tag{A7}$$

$$= \overline{W} + \lambda \left(N\gamma\tilde{\delta} + \gamma \sum_i \tilde{\varepsilon_i} + z \right), \tag{A8}$$

where λ is the sensitivity of the asset price to order flows. Also, the profit of informed trader i is expressed as

$$\pi_i = (\tilde{W} - P)x_i. \tag{A9}$$

In this setting, we can solve for the equilibrium that satisfies the two conditions: profit maximization by the informed and market efficiency.

In Equations (A7) and (A8), λ is an illiquidity measure and its inverse, $1/\lambda$, is sometimes called the "depth" of the market. If a market is very liquid, one would expect that combined trades, ω, may not affect the asset price very much, and hence the level of λ is low. Our goal is to solve λ as a measure of illiquidity from the equilibrium conditions.

First, informed trader i's problem is

$$Max: E[\pi_i \mid \tilde{\delta} + \tilde{\varepsilon_i}] = E\left[\left\{ \overline{W} + \tilde{\delta} - \overline{W} - \lambda \left(x_i + (N-1)\gamma\tilde{\delta} + \gamma \sum_{j \neq i} \tilde{\varepsilon}_j \} + z \right) \right\} x_i \mid \tilde{\delta} + \tilde{\varepsilon_i} \right]$$

$$= \{E[\tilde{\delta} \mid \tilde{\delta} + \tilde{\varepsilon_i}] - \lambda x_i - \lambda(N-1)\gamma E[\tilde{\delta} \mid \tilde{\delta} + \tilde{\varepsilon_i}]\}x_i$$

$$= -\lambda x_i^2 + x_i \{1 - \lambda(N-1)\gamma\} E\left[\tilde{\delta} \mid \tilde{\delta} + \tilde{\varepsilon_i}\right]. \tag{A10}$$

[27] Equation (A6) comes from the property of the multivariate normal distribution. Let two random variables, X_1 and X_2, be jointly normally distributed so that $\binom{X_1}{X_2} \sim N[\binom{\mu_1}{\mu_2}, \begin{pmatrix}\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}]$. Then, we can show that $E[X_1|X_2] = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2)$, and $Var[X_1|X_2] = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. For details, see Anderson (1984).

The first-order condition of Equation (A10) gives $-2\lambda x_i + \{1 - \lambda(N-1)\gamma\} E\left[\tilde{\delta} \mid \tilde{\delta} + \tilde{\varepsilon_i}\right] = 0$. Thus,

$$\begin{aligned} x_i &= \frac{1}{2\lambda}\{1 - \lambda(N-1)\gamma\} E[\tilde{\delta} \mid \tilde{\delta} + \tilde{\varepsilon_i}] \\ &= \frac{1}{2\lambda}\{1 - \lambda(N-1)\gamma\} \frac{v_\delta}{v_\delta + v_\varepsilon}(\tilde{\delta} + \tilde{\varepsilon_i}). \end{aligned} \quad \text{(A11)}$$

Therefore, from Equations (A2) and (A11), we have $\gamma = \frac{1}{2\lambda}\{1 - \lambda(N-1)\gamma\} \frac{v_\delta}{v_\delta + v_\varepsilon}$, which in turn leads to

$$\gamma = \frac{\frac{v_\delta}{v_\delta + v_\varepsilon}}{\lambda\left\{2 + \frac{v_\delta}{v_\delta + v_\varepsilon}(N-1)\right\}}. \quad \text{(A12)}$$

Next, from Equations (A6) and (A8), the market efficiency condition is equivalent to

$$\begin{aligned} \lambda &= \frac{Cov\left[\tilde{\delta}, N\gamma\tilde{\delta} + \gamma\sum_i \tilde{\varepsilon_i} + z\right]}{Var\left[N\gamma\tilde{\delta} + \gamma\sum_i \tilde{\varepsilon_i} + z\right]} \\ &= \frac{N\gamma v_\delta}{N^2\gamma^2 v_\delta + \gamma^2 N v_\varepsilon + v_z} \\ &= \frac{N v_\delta}{\gamma(N^2 v_\delta + N v_\varepsilon) + \frac{1}{\gamma} v_z}. \end{aligned} \quad \text{(A13)}$$

Plugging Equation (A12) into Equation (A13) gives

$$\lambda = \frac{v_\delta}{(N+1)v_\delta + 2v_\varepsilon}\sqrt{\frac{N(v_\delta + v_\varepsilon)}{v_z}}. \quad \text{(A14)}$$

Note that initially we assumed informed traders observe a signal with noise, $\tilde{\varepsilon_i}$ $(i = 1, 2, 3, \ldots, N)$. Now suppose there is no noise in the signal so that $v_\varepsilon = 0$. Then, Equation (A14) is reduced to

$$\lambda = \frac{\sqrt{N v_\delta}}{(N+1)\sqrt{v_z}}. \quad \text{(A15)}$$

In this study, Equation (A14) and Equation (A15) are used as the primary basis of our two illiquidity measures.

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* 17:223–49.

Anderson, T. 1984. *An Introduction to Multivariate Statistical Analysis,* 2nd ed. New York: John Wiley & Sons.

Avramov, D., and T. Chordia. 2006. Asset Pricing Models and Financial Market Anomalies. *Review of Financial Studies* 19:1002–40.

Baker, M., and J. Stein. 2004. Market Liquidity as a Sentiment Indicator. *Journal of Financial Markets* 7:271–99.

Bessembinder, H., and I. Kalcheva. 2006. Liquidity Biases in Asset Pricing Tests. Working Paper, University of Utah.

Black, F., M. Jensen, and M. Scholes. 1972. The Capital Asset Pricing Model: Some Empirical Tests, in M. Jensen (ed.), *Studies in the Theory of Capital Markets*. New York: Praeger Publishers.

Blume, M., and R. Stambaugh. 1983. Biases in Computed Returns: An Application to the Size Effect. *Journal of Financial Economics* 12:387–404.

Brennan, M., T. Chordia, and A. Subrahmanyam. 1996. Cross-Sectional Determinants of Expected Returns, in B. Lehmann (ed.), *In Honor of Fischer Black*. Cary, NC: Oxford University Press.

Brennan, M., T. Chordia, and A. Subrahmanyam. 1998. Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns. *Journal of Financial Economics* 49:345–73.

Brennan, M., N. Jegadeesh, and B. Swaminathan. 1993. Investment Analysis and the Adjustment of Stock Prices to Common Information. *Review of Financial Studies* 6:799–824.

Brennan, M., and A. Subrahmanyam. 1995. Investment Analysis and Price Formation in Securities Markets. *Journal of Financial Economics* 38:361–81.

Brennan, M., and A. Subrahmanyam. 1996. Market Microstructure and Asset Pricing: On the Compensation for Illiquidity in Stock Returns. *Journal of Financial Economics* 41:441–64.

Brennan, M., and A. Wang. 2006. Asset Pricing and Mispricing. Working Paper, University of California at Los Angeles.

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* 65:111–30.

Chordia, T., S. Huh, and A. Subrahmanyam. 2007. The Cross-Section of Expected Trading Activity. *Review of Financial Studies* 20:709–740.

Connor, G. 1984. A Unified Beta Pricing Theory. *Journal of Economic Theory* 34:13–31.

Eisfeldt, A. 2004. Endogenous Liquidity in Asset Markets. *Journal of Finance* 59:1–30.

Eleswarapu, V., and M. Reinganum. 1993. The Seasonal Behavior of the Liquidity Premium in Asset Pricing. *Journal of Financial Economics* 34:373–86.

Fama, E., and K. French. 1992. The Cross-Section of Expected Stock Returns. *Journal of Finance* 47:427–66.

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* 81:607–36.

Gallant, R., P. Rossi, and G. Tauchen. 1992. Stock Prices and Volume. *Review of Financial Studies* 5:199–242.

Gibbons, M., S. Ross, and J. Shanken. 1989. A Test of the Efficiency of a Given Portfolio. *Econometrica* 57:1121–52.

Glosten, L., and L. Harris. 1988. Estimating the Components of the Bid-Ask Spread. *Journal of Financial Economics* 21:123–42.

Green, C. 2006. The Value of Client Access to Analyst Recommendations. *Journal of Financial and Quantitative Analysis* 41:1–24.

Hasbrouck, J. 1999. The Dynamics of Discrete Bid and Ask Quotes. *Journal of Finance* 54:2109–42.

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.

Hvidkjaer, S. 2006. A Trade-Based Analysis of Momentum. *Review of Financial Studies* 19:457–91.

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.

Jegadeesh, N., and S. Titman. 2001. Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. *Journal of Finance* 56:699–720.

Johnson, T. 2005. Dynamic Liquidity in Endowment Economies. *Journal of Financial Economics* 80:531–62.

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* 53:1315–35.

Kumar, P., and D. Seppi. 1994. Information and Index Arbitrage. *Journal of Business* 67:481–509.

Lee, C., and M. Ready. 1991. Inferring Trade Direction from Intraday Data. *Journal of Finance* 46:733–47.

Merton, R. 1980. On Estimating the Expected Rate of Return on the Market. *Journal of Financial Economics* 8:323–62.

Newey, W., and K. West. 1987. A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55:703–8.

Newey, W., and K. West. 1994. Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies* 61:631–53.

Pástor, L., and R. Stambaugh. 2003. Liquidity Risk and Expected Stock Returns. *Journal of Political Economy* 113:642–85.

Roll, R. 1984. A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *Journal of Finance* 39:1127–39.

Sadka, R. 2006. Momentum and Post-Earnings-Announcement Drift Anomalies: The Role of Liquidity Risk. *Journal of Financial Economics* 80:309–49.

Shanken, J. 1985. Multibeta CAPM or Equilibrium APT? A Reply. *Journal of Finance* 40:1189–96.

Shanken, J. 1987. Multivariate Proxies and Arbitrage Pricing Relations: Living with the Roll Critique. *Journal of Financial Economics* 18:91–110.

Shumway, T. 1997. The Delisting Bias in CRSP Data. *Journal of Finance* 52:327–40.

Spiegel, M., and X. Wang. 2005. Cross-Sectional Variation in Stock Returns: Liquidity and Idiosyncratic Risk. Working Paper, Yale University.

Subrahmanyam, A. 1991. A Theory of Trading in Stock Index Futures. *Review of Financial Studies* 4:17–51.