July 10, 2006

Theory-Based Illiquidity and Asset Pricing

Tarun Chordia*

Sahn-Wook Huh**

Avanidhar Subrahmanyam***

*Goizueta Business School, Emory University, Atlanta, GA 30327.

**Faculty of Business, Brock University, St. Catharines, Ontario, Canada L2S 3A1.

***The Anderson School, 110 Westwood Plaza, University of California at Los Angeles, Los Angeles, CA 90095-1481.

Abstract

Theory-Based Illiquidity and Asset Pricing

Many proxies of illiquidity have been used in the literature that relates illiquidity to asset prices. These proxies generally are motivated from an empirical standpoint. However, theory-based formulae for an illiquidity measure, namely, the Kyle lambda, are available in closed-form and empirical proxies can be used to estimate such lambdas. We estimate such lambdas for a comprehensive sample of NYSE and Nasdaq stocks. There is convincing evidence that these theory-based estimates of illiquidity are priced in the cross-section of expected stock returns.

The question of whether investors demand a higher return from less liquid securities is an enduring one in financial economics. In the seminal paper on this issue, Amihud and Mendelson (AM) (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 role of liquidity as a determinant of expected returns. Further, 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 that of how to measure illiquidity. Other than direct empirical measurements of illiquidity by quantities such as the quoted bid-ask spread, the approach taken in the literature has been to propose empirical arguments and econometric techniques to measure illiquidity. For example, Amihud (2002) proposes the ratio of absolute return to volume as a measure of illiquidity, and Brennan and Subrahmanyam (1996), based on the analysis of Glosten and Harris (1988) propose to measure illiquidity by the relation between price changes and order flow. Pástor and Stambaugh (2003) measure illiquidity by the extent to which returns reverse upon high volume, 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, and we will not duplicate his efforts by describing these measures in detail.

These empirical proxies have added considerable to our understanding of illiquidity.

[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.

However, there are some issues related to this literature that are matters of concern. First, the mixed results that have been obtained using some of these measures have made the state of the literature hard to interpret. Thus, for example, Brennan and Subrahmanyam (1996) find a negative relation between bid-ask spread and expected returns which contradicts the liquidity premium argument. Spiegel and Wang (2005) do not find a significant relation between expected returns and a variety of empirically-motivated illiquidity proxies, which also appears to muddy the conclusions on whether illiquidity is related to asset returns. A related problem is that the empirical arguments proposed to justify such proxies have often been ad hoc from a theoretical standpoint. For example, while many microstructure theories have been developed, no such theory indicates that illiquidity can be captured by the Amihud (2002) construct of the ratio of absolute return to volume. A third issue is that illiquidity is endogenous and depends on many variables that are related to asset prices via other models. For example, illiquidity depends on volatility, but volatility is related to expected returns via traditional risk-return arguments.

In this paper, we propose an alternative approach to measure illiquidity and relate it to expected asset returns. Specifically, we turn to theory in order to consider illiquidity estimates that can be directly estimated from closed-form expressions from an economic model. 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, obtains from the Kyle (1985) model and its adaptation by Admati and Pfleiderer (1988) to explain intraday patterns. The advantage of using 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 cost to a researcher.

We estimate two variants of closed-form expressions for Kyle lambdas using a comprehensive set of NYSE and Nasdaq stocks over a long time-period and examine whether they are priced in the cross-section of stock returns. The empirical proxies for the inputs to the Kyle (1985) model are along the lines of those used by Brennan and Subrahmanyam (1995), supplemented by other plausible arguments. We examine the time-series behavior of such lambdas and find a decline in some of the measures over time, which mirrors the behavior of other illiquidity proxies such as bid-ask spreads (Jones, 2002). We also find convincing evidence that Kyle lambdas are priced in the cross-section of stock returns, after controlling for known characteristics such as book/market and momentum and known sources of risk such as the Fama and French (1993) factors.

1 Estimates of Kyle Lambdas

In this section, we provide the theoretical background for our lambda estimation. We estimate two versions of Kyle lambdas, with and without signal noise. We begin by linking illiquidity and asset pricing in the context of Kyle lambdas.

1.1 *A Simple Link between Kyle Lambdas and Illiquidity Pricing*

The Kyle (1985) model does not provide a direct link to illiquidity pricing. However, this can be incorporated into the framework assuming that the liquidity trader is the marginal agent. Thus, suppose an asset is traded over two periods. In period 2, it pays off $\tilde{W} = \bar{W} + \delta$ where δ is a normally distributed variable with a mean of zero. A liquidity trader's demand is denoted by D. There is also a non-discretionary liquidity demand, which again is normally distributed with zero mean, and equals $z - D$. For a given λ, the price impact of his trade is given by $E[(\xi - F)D] = \lambda D^2$.

We now assume that the asset's supply is normalized to one share, and at date 0 (prior to trade at date 1), the liquidity trader contemplates investing in the stock.[2] We assume that the risk-free discount rate is zero across dates 0 and 2 for convenience. 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

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

Thus, the expected price change across dates 0 and 2 is given by[3] λD^2, and is thus proportional to λ. Thus, expected future returns are linearly related to λ.

Our study uses structural estimates of two versions of Kyle lambdas, with and without signal noise. We now present the theory-based illiquidity measures and discuss how to estimate them using proxy variables.

1.2 *An Illiquidity Measure without Noise in the Signal*

When the informed traders observe, without any noise, a signal that is informative about the liquidation value of an asset, Appendix shows in detail that an illiquidity measure, λ, 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 payoff innovations, and v_z is the variance of uninformed trades. Dividing both sides of Eq.(1) by the price P in

[2]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.

[3]Note that the expected price change across dates 1 and 2 is zero (the date 1 price is semi-strong efficient) and the expected price change across dates 0 and 1 is equal to the expected price change across dates 0 and 2. Thus, it the expected price change across dates 0 and 2 is the only unique, non-zero expected price change in our model.

order to get a price-scaled illiquidity measure, we have[4]

$$\begin{aligned}\frac{\lambda}{P} &= P^{-1}\frac{\sqrt{Nv_\delta}}{(N+1)\sqrt{v_z}} \\ &= \frac{\sqrt{N}\sqrt{Var(R)}}{(N+1)\sqrt{v_z}} \\ &= \frac{N^{0.5}std(R)}{(N+1)std(z)}, \qquad (2)\end{aligned}$$

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

To estimate ILLIQ_1 for each stock in each month, we employ proxy variables as inputs for each of the original variables in Eq.(2) as follows:

N: One plus the number of analysts following a firm in each month, notated as ANA.[5]

$std(R)$: The standard deviation of daily returns in the previous month (month $t-1$), notated as STD(RET). To obtain this variable for each month, we use firms that have at least 10 daily returns in the previous month from the CRSP daily file.

$std(z)$: The average of daily dollar volume (in \$million) in the previous month, notated as AVG(DVOL).[6] To obtain this variable for each month, we use firms that have at least 10 daily trading records in the previous month from the CRSP daily file.

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

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

[6]Note that because 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.

1.3 *An Illiquidity Measure with Noise in the Signal*

When the informed traders observe diverse signals with independent and identically distributed error terms, the Appendix shows that the illiquidity 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 payoff innovations, v_z is the variance of uninformed trades, and v_ε is the variance of signal innovations. Dividing both sides of Eq.(3) by P_{t-1}, we have

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

Eq.(4) is our second measure of illiquidity used in this study, and we call it ILLIQ_2.

To estimate ILLIQ_2 for each stock in each month, we employ proxy variables for each of the original variables in Eq.(4) as follows:

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 EPSs from the most recent eight quarters.

v_ε: This variable is proxied by ESURP-sqr, which is the squared value of earnings surprise (ESURP), where ESURP is the absolute value of the current earnings per share (EPS) minus the EPS four quarters ago.

v_z: We proxy this by AVG(DVOL)-sqr, which is the squared value of the average of daily dollar volume (in $million) in the previous month. To obtain this variable for each month, we use firms that have at least 10 daily trading records in the previous month from the CRSP daily file.

1.4 *Estimation of the Illiquidity Measures*

To estimate our illiquidity measures (ILLIQ_1 and ILLIQ_2) according to Eq.(2) and Eq.(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. 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 in for NYSE/AMEX stocks are 1,845.1 and 1,683.1, respectively. Those for NASDAQ stocks are 2,663.8 and 1,967.0, respectively.

Table I contains the descriptive statistics for the input variables of the first illiquidity measure, ILLIQ_1. As one would expect, the mean of ANA for NASDAQ (interchangeably, the "OTC market") stocks (3.15) is much lower than that for NYSE/AMEX (interchangeably, the "exchange market") stocks (5.52). Given that firm size in the NASDAQ market is often smaller, average daily dollar volume, AVG(DVOL), in this market ($3.61 million) is lower than that in the exchange market ($5.26 million). Because NASDAQ stocks are more high-tech oriented as well as smaller, it is not surprising to see that daily return volatility, STD(RET), in the OTC market (4.4%) is far higher than that in the exchange market (2.7%).

To check the descriptive statistics of the input variables for the second illiquidity measure, ILLIQ_2, we again see that ANA and AVG(DVOL)-sqr are qualitatively similar to the corresponding input variables for ILLIQ_1. In the same context, the price level in the exchange market ($31.31) is much higher than in the OTC market ($14.16). While earnings surprise variables (ESURP and ESURP-sqr) are higher in the exchange market,

earnings volatility variables (EVOLA and EVOLA-sqr) are way higher in the NASDAQ market. Also note that ESURP-sqr is much more fluctuating across firms in the NASDAQ market.

2 Methodology

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 return on the k-th factor ($k = 1, 2, ..., L$) at time t. The exact or equilibrium version of the arbitrage pricing theory (APT) in which the market portfolio is well diversified with respect to the factors (Connor, 1984; Shanken, 1985, 1987) implies that the expected excess returns may be written as

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

where R_{Ft} is the return on the risk-free asset and θ_{kt} is the risk premium on the factor portfolio k. Plugging Eq.(6) into Eq.(7), 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 and return on the factor k.

Our goal is to test whether the two illiquidity measures theoretically derived in Section 1 from the strategic microstructure model have incremental explanatory power for returns relative to the 3-factor FF 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:

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

where $ILLIQ\_i_{jt}$ ($i = 1$ 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 ($m = 1$, ..., M) for security j in month t. Note here that the right-hand side variables in Eq.(8) are all one-period preceding in order to ensure we capture pure predictive relations. Under the null hypothesis that expected excess returns depend only on the risk characteristics of the returns notated by β_{jk}, then ϕ and c_m ($m = 1$, ..., M) will be zero. This hypothesis can be tested in principle by estimating the factor loadings each month using the past data, conducting a cross-sectional regression for each month in which the independent variables are an illiquidity measure, factor loadings, and other non-risk characteristics, and then averaging the monthly coefficients over time and computing their standard errors. This basic Fama-MacBeth approach, however, will present a problem if the factor loadings are measured with errors. One method dealing with this measurement error problem is to use the information from the first-stage regressions (in which the factor loadings are estimated) to correct the efficient estimates in the second-stage regressions.[7] Our approach to correct the bias, however, does not rely on information taken from the first-stage regressions.

In order to correct the bias, we perform risk adjustments in returns using the Fama-French (FF, 1993) 3 factors (MKT_t, SMB_t, and HML_t)[8] 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

[7]For this approach, see Litzenberger and Ramaswamy (1979) and Lehmann (1990).

[8]MKT is the excess return on the market portfolio, SMB is the return on a zero net investment portfolio which is long in small firms and short in large firms, and HML is the return on a zero net investment portfolio which is long in high book-to-market firms and short in low book-to-market firms.

intercept and the residual, i.e.,

$$\begin{aligned}\tilde{R}_{jt}^{*1} &= (\tilde{R}_{jt} - R_{Ft}) - (\widehat{\beta}_{j1}^{*} MKT_t + \widehat{\beta}_{j2}^{*} SMB_t + \widehat{\beta}_{j3}^{*} HML_t) \\ &= \widehat{\alpha}^{*} + \widehat{\tilde{e}}_{jt}^{*}, \end{aligned} \tag{9}$$

after conducting regressions in Eq.(7) (but with a constant term α) using the *whole* sample range (from January 1972 to December 2002 for NYSE/AMEX stocks and from January 1983 to December 2002 for NASDAQ stocks) of the data.[9] We call this risk-adjusted return ($\tilde{R}_{jt}^{*1}$) as *FF3-adj EXSRET1*. In the second method, we first estimate the factor loadings, β_{jk}, *each month* over the sample period for all securities using the time series of the past 60 months (at least 24 months) with Eq.(7). Given the current month 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 then can compute the risk-adjusted return on each of the securities, $\tilde{R}_{jt}^{*2}$, for each month t as follows:

$$\tilde{R}_{jt}^{*2} = (\tilde{R}_{jt} - R_{Ft}) - (\widehat{\beta}_{j1}^{**} MKT_t + \widehat{\beta}_{j2}^{**} SMB_t + \widehat{\beta}_{j3}^{**} HML_t). \tag{10}$$

We call this risk-adjusted return ($\tilde{R}_{jt}^{*2}$) as *FF3-adj EXSRET2*.

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

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

Note that the error term in Eq.(11) is different from that in Eq.(8), because the error in Eq.(11) also contains terms arising from the measurement error associated with the factor loadings.

[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.

To check if illiquidity is priced, we report, as in Brennan et al. (1996, 1998), two types of statistics based on regressions in Eq.(11), in addition to the statistics based on regressions with the dependent variable in Eq.(11) being risk-*unadjusted* excess returns (we call this unadjusted return as EXSRET). The first is the standard Fama-MacBeth (1973) estimator (notated as "Raw" in Tables V and VI), and the second is the constant term from the OLS regression of the time series of the coefficient estimates from Eq.(11) on the factor portfolio returns, which is referred to as the "purged" estimator (notated as "Purged" in Tables V and VI). For our purposes, we estimate the vector of coefficients $\mathbf{c}_t = [c_{0t}\ \phi_t\ c_{1t}\ c_{2t}...c_{Mt}]'$ from Eq.(11) each month with a simple OLS regression as

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

where $i = 1$ or 2, $\mathbf{Z}_t = [ILLIQ\_i\ \ Z_1\ Z_2...Z_M]'$, and $\tilde{\mathbf{R}}_{t+1}^{*i}$ is the vector of risk-adjusted excess returns based on Eq.(9) or Eq.(10). The standard Fama-MacBeth (1973) estimator is the time-series average of the monthly coefficients, and the standard error of this estimator is taken from the time series of monthly coefficient estimates, $\widehat{\mathbf{c}}_t$. Note that although factor loadings are estimated with error in Eq.(7), this error affects only the dependent variable, $\tilde{R}_{t+1}^{*i}$ as we see in Eq.(9), Eq.(10), and Eq.(11). While the factor loadings will be correlated with vector $\mathbf{Z}_t = [ILLIQ\_i\ \ Z_1\ Z_2...Z_M]'$, there is no *a priori* reason to believe that the errors in the estimated loadings will be correlated with the vector $\mathbf{Z}_t$. This implies that the coefficient vector $\widehat{\mathbf{c}}_t$ estimated in Eq.(11) is unbiased.

However, if the errors in the estimated factor loadings are correlated with the explanatory variables $\mathbf{Z}_t = [ILLIQ\_i\ \ Z_1\ Z_2...Z_M]'$, 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 obtain

a purged estimator for each of the explanatory variables: i.e., the constant term (and its t-value) from the regression of the monthly coefficients ($\widehat{\mathbf{c}}_t$) estimated in Eq.(11) on the time series of FF 3 factor realizations. This estimator, which was developed by Black et al. (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, and Descriptive Statistics

For this study, we use data at a monthly frequency over the 372 months (31 years: 197201-200212) for NYSE/AMEX stocks and the 240 months (20 years: 198301-200212) for NASDAQ stocks. In some cases where accounting variables and other data are available only on a yearly (or quarterly) basis, we keep those values constant for 12 months (or 3 months) in the regressions.[10]

The three dependent variables (EXSRET, FF3-adj EXSRET1, and FF3-adj EXSRET2) defined in Section 2 for the Fama-MacBeth (1973) regressions are obtained or estimated using the CRSP monthly file and the FF 3 factors are available from Kenneth French's web site. In addition to the variables mentioned above, we use six firm 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).

SIZE: The natural logarithm of MV.

[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. Those available only on a quarterly basis are accounting performance-related variables: ESURP-sqr and EVOLA-sqr.

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).

BM_Trim: The trimmed 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 3 months (from month t-1 to month t-3).

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

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

MOM4: The compounded holding period return over the 3 months from month t-9 to month t-12. For each of the above 4 momentum variables to exist, a stock should have all three consecutive monthly returns over the corresponding 3-month period.

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 average number of component stocks used each month in the Fama-MacBeth (1973)-type cross-sectional regressions for NYSE/AMEX stocks is 1,845.1, while that for NASDAQ stocks is 2,663.8.

Table II reports the time-series average values of monthly means, medians, standard deviations (STD), and other descriptive statistics for our key variables. The values of each statistic are first computed cross-sectionally then averaged in the time-series over the sample period. Noticeable is that the average means of the two illiquidity measures, ILLIQ_1 and ILLIQ_2, in the NASDAQ market are higher by the factor of almost 5 than in the exchange market. The variations of the measures across stocks in the NASDAQ market are also far higher. Considering that the NASDAQ market is characterized by young, small, and high-tech firms on average, these results seem reasonable. This points out that proxying (il)liquidity by share turnover only may be misleading. Chordia, Huh, and Subrahmanyam (2005) document that turnover in the Nasdaq market has been much higher than that in the exchange market. However, this may be partly because of double counting in the NASDAQ market (Atkins and Dyl, 1997), indicating that despite the higher turnover in the NASDAQ market we cannot safely say that the NASDAQ market is more liquid than the exchange market.

To visually examine the trends of the illiquidity measures, we plot the value-weighted illiquidity measures in Figures 1 and 2 for the two markets over the sample period. As Figure 1(a) shows, value-weighted ILLIQ_1 of NYSE/AMEX stocks reaches a peak around 1975 and then exhibits a decreasing time trend, suggesting that market liquidity has improved since the mid-1970s. As mentioned above, the average level of monthly illiquidity in the OTC market [Figure 1(b)] is consistently higher, but the measure is also decreasing from the early 1990s. Reflecting the recent economic recession, ILLIQ_1 in the OTC market shows a more salient increase in 2001-2002 than that in the exchange market. As can be seen in Figure 2, the trend of ILLIQ_2 is qualitatively similar to that of ILLIQ_1, with its volatility and 2001-2002 level increase being more pronounced in both markets.

Other discernible facts are as follows. NYSE/AMEX stocks are larger by a big difference, and NASDAQ stocks are more likely to be growth stocks (see BM_Raw). The 4 momentum variables (MOM1-MOM4) are consistently negative in the NASDAQ market, while they are positive in the exchange market. In addition, the illiquidity measures, firm size (MV), and the book-to-market ratio (BM_Raw) tend to be right-skewed.

Next, we examine the average correlation coefficients between our explanatory variables in Table III.[11] The lower and upper triangles in the table present the correlations for NYSE/AMEX and NASDAQ stocks, respectively. Our two illiquidity measures are highly correlated each other in both markets: 64% in the exchange market, and 46% in the NASDAQ market. The two measures are negatively correlated (and mostly significant) with the 4 momentum variables, suggesting that good past price performance of a stock contributes to improving the liquidity of that stock. Given that institutional investors, as dominant participants in stock markets, prefer to trade large stocks, it is not surprising to observe that correlation of SIZE with the two illiquidity measures is negative and statistically significant at any conventional level in both markets. SIZE is also positively correlated with the 4 momentum variables. An interesting feature is that the correlation coefficients between the book-to-market ratio and the two illiquidity measures are positive and statistically significant at the 1% level in both markets. This indicates that value stocks are likely to be more illiquid. The 4 momentum variables tend to be positively correlated each other, but correlation is higher between any two momentum variables with non-neighboring periods.

[11] For brevity, 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 highly correlated (coefficients greater than 94%) in both markets.

4 Empirical Results

4.1 *Features of the Portfolios Formed on Illiquidity and Size*

Before moving on to regression analyses, we report the average values of monthly return, firm size, and illiquidity for the 25 portfolios formed on illiquidity and firm size. For this purpose, each month we first sort sample stocks by ILLIQ_1 in an ascending order and split into 5 portfolios with the equal number of stocks. Then, each of the 5 portfolios is again sorted by firm size (MV) and split into 5 portfolios, resulting in the 25 portfolios. Next, the average values of return, size, and illiquidity are computed each month for each of the 25 portfolios, and the time-series averages of the 3 variables over the sample period are reported in each panel of Table IV.

Panel A in Table IV shows that for a given size group (especially size groups 1-3) the average return tends to increase with illiquidity in both markets, while for a given illiquidity group the average return tends to decrease with firm size. The former indicates that theory-based illiquidity is priced, and the latter confirms the small-firm effect evidenced by many researchers. The t-values (italicized in Panel A) also demonstrate that monthly portfolio returns are mostly significantly different from zero. In particular, note that the average return of the portfolio with smallest size and highest illiquidity is 3 times higher than that of the other extreme portfolio (with largest size and lowest illiquidity) in the NYSE/AMEX market, and 8 times higher in the NASDAQ market. Within a given size group, the return difference between the two extreme portfolios (the lowest illiquidity group vs. the highest illiquidity group) is not so pronounced as that between the two extreme portfolios above, but the paired-sample t-test (where the null hypothesis is that the return difference is zero) shows t-values of 3.58, 1.97, and 1.66 for the first 3 size groups, respectively, for NYSE/AMEX stocks, while they are 8.25, 2.70,

and 1.88, respectively, for NASDAQ stocks. As can be seen in Panels B and C, average firm size (within a given size group) is negatively related to illiquidity, and average illiquidity (within a given ILLIQ_1 group) is mostly negatively related to firm size. To save space, we do not report the analog of a table with ILLIQ_2, but the result is qualitatively similar to the result with ILLIQ_1.

4.2 *Cross-Sectional Regressions*

We have observed in Table IV that within a given size group the average portfolio return is likely to increase with illiquidity, suggesting that theory-based illiquidity is priced. In this section, we formally test if our two illiquidity measures predict returns. As mentioned above, our test involves the following cross-sectional regression estimated at the monthly frequency:

$$\tilde{R}^*_{jt+1} = c_{0t} + \phi_t ILLIQ\_i_{jt} + \sum_{m=1}^{M} c_{mt} Z_{mjt} + \tilde{e}'_{jt+1}, \text{i} = 1\text{or}2, \tag{12}$$

where $\tilde{R}^*_{jt+1}$ represents one of the risk-unadjusted excess return (EXSRET) and two risk-adjusted excess returns (FF3-adj EXSRET1 and FF3-adj EXSRET2) defined and estimated in Section 2, $ILLIQ\_i_{jt}$ is either of our two theory-based illiquidity measures (ILLIQ_1 and ILLIQ_2) derived and estimated in Section 1, and Z_{mjt} denotes firm characteristic m for stock j in month t.[12]

Considering that, as in Fama and MacBeth (1973), $\tilde{R}^*_{jt+1}$ is the monthly stock return considered to be an *i.i.d.* process, we report the standard Fama-MacBeth statistics (the time-series average of the estimated coefficients from the equation above and its t-statistic) for the columns under "Unadjusted" and "Raw" in Tables V and VI. Following Black et al. (1972), however, we also report a purged estimator (the columns under

[12]Note that unlike the contemporaneous regressions in Fama and MacBeth (1973), our explanatory variables are lagged one period.

"Purged") as a check on the robustness of our results by estimating the constant term (and its t-value) from the regression of the monthly coefficients estimated in Eq.(12) on the time series of FF 3 factor realizations.

Along with the average coefficients and t-statistics, we provide 2 other types of statistics in the tables: the average of the adjusted R-squared 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). Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

The regression results with ILLIQ_1 in Eq.(12) are presented in Table V, while those with ILLIQ_2 are in Table VI. As we see in Table V, the average number of component stocks used in the monthly regressions for NYSE/AMEX stocks is 1797.7-1,845.1, and that for NASDAQ stocks is 2406.3-2,663.8, depending on data availability in the variables. *Avg R-sqr* is in the 3-5% range in the exchange market, and that for NASDAQ stocks is slightly lower. The explanatory power of the regressions tends to be higher with the unadjusted excess returns (EXSRET) than with the risk-adjusted returns (FF3-adj EXSRET1, FF3-adj EXSRET2) in both markets. Given that ILLIQ_2 requires more input variables, Table VI exhibits that the average number of component stocks used in the regressions with this illiquidity measure decreases by 7-8% to 1667.8-1683.1 for NYSE/AMEX stocks, and by 21-26% to 1907.3-1967.0 for NASDAQ stocks. Other aspects about the explanatory power are similar to those with ILLIQ_1.

We first discuss the results from the Fama-MacBeth regressions of EXSRET on ILLIQ_1 as well as other firm characteristics that are best known to be associated with expected returns, namely, SIZE, BTM, and the 4 momentum variables (MOM1-MOM4). Table V shows that the average coefficients of ILLIQ_1 are positive and statistically sig-

nificant at the 1% level in both markets after controlling for other conventional firm characteristics, confirming the hypothesis that stocks with higher illiquidity are expected to have higher (excess) returns. The coefficients of SIZE and BTM are respectively negative and positive, and both are statistically significant, which is consistent with previous studies such as Fama and French (1992). We also find that the sensitivity of these variables to returns are much greater in the NASDAQ market than in the exchange market. The 4 momentum variables are all strongly positively related to returns in the exchange market, while they become weaker in the OTC market.

We now consider whether the relations observed above are maintained when the dependent variable is risk-adjusted using the FF 3 factors. The raw and purged estimates of illiquidity and characteristic rewards ($\widehat{\phi}$, $\widehat{c}_m$) for returns adjusted by the first method in Section 2 (FF3-adj EXSRET1) are presented in the next two columns of Table V. By risk-adjusting, the coefficients tend to attenuate slightly, but the relations are essentially unchanged, with the levels of statistical significance becoming even reinforced in many cases. ILLIQ_1 continues to be strongly positively related to risk-adjusted returns, firm size is negatively related to returns, and a higher book-to-market ratio predicts higher returns in both markets. Overall, SIZE and BTM play more important roles in the NASDAQ market than in the NYSE/AMEX market in predicting stock returns. The momentum variables again demonstrate that better price performance in the past is expected to provide higher returns in the current month, especially in the NYSE/AMEX market. This confirms the continuation of short-term returns documented by Jegadeesh and Titman (1993) and Fama and French (1996). In addition, there is little difference between the raw and purged estimates as one would expect if the estimation errors in factor loadings are uncorrelated with the vector of explanatory variables.

In the remaining two columns, we report the raw and purged estimates of illiquidity

and characteristic rewards ($\widehat{\phi}$, $\widehat{c}_m$) for excess returns (FF3-adj EXSRET2) which are now risk-adjusted by the second method in Section 2. First, the impact of ILLIQ_1 on risk-adjusted returns is virtually the same as the result with FF3-adj EXSRET1: positive and statistically significant in the two markets. Although SIZE and BTM continue to have strong impact on excess returns in the NASDAQ market, the statistical significance of the two variables is now weaker in the exchange market. In particular, the purged estimates for these variables exhibit that their influence on returns are only marginal. Another discernible feature is that the momentum variables now show a contrast between the two different markets. Using FF3-adj EXSRET2, these variables are strongly positively related to returns in the exchange market, whether the estimates are raw or purged. In the NASDAQ market, however, the coefficients of the variables, whether they are raw or purged, become insignificant without exception. This is an interesting aside which deserves investigation in future research.

Next, we investigate in Table VI how the effects of illiquidity and other firm characteristics on returns change when we employ the second measure, ILLIQ_2, in the regressions. Most noticeable is that the coefficients of ILLIQ_2 are not statistically different from zero in the NASDAQ market, whichever returns are used as a dependent variable. For NYSE/AMEX stocks, however, we still find a significant relation between the required rate of return and illiquidity after accounting for the effects of other characteristics. It is possible that the analyst forecasts for Nasdaq stocks are more prone to error, which may, in turn, induce noise in the estimates of ILLIQ_2, which relies on the analyst forecast to estimate signal noise. As observed in Table V about the roles of SIZE and BTM, the magnitude of the coefficients and the levels of their statistical significance are again larger in the NASDAQ market than in the exchange market. The momentum effects in the NYSE/AMEX market are similar to those in Table V, but the effects are more

salient in the NASDAQ market compared to the previous result. That is, the coefficient estimates of more recent past returns (MOM1-MOM2) for NASDAQ stocks are larger by 40-150% than those in Table V and are now statistically significant at the 5% level even for the purged estimates in the regression with FF3-adj EXSRET2.

4.3 *Robustness Checks*

As a check on the robustness of our results, we have already used three different types of excess returns (EXSRET, FF3-adj EXSRET1, and FF3-adj EXSRET2), further reporting the purged estimates of Black et al. (1972) as well as the standard Fama-MacBeth (1973) estimates. Despite the computational and programming burden, we have also considered the effects of our choices in input variables used to estimate the two illiquidity measures on the results. As explained in Section 1, the two important input variables in estimating ILLIQ_1 and ILLIQ_2 are, among others, the standard deviation of daily returns in month $t-1$ and the average of daily dollar volume in month $t-1$. Other than these, we have obtained standard deviations of returns computed each month with daily returns in month $t-2$, daily returns in the past 36 months, or monthly returns in the past 60 months. Moreover, as a proxy for $std(R)$, the idiosyncratic risk against the FF 3 factors has been computed each month using the data from the past 60 months in line with Spiegel and Wang (2005). For the average volume as a proxy for $std(z)$, many candidates have also been computed each month with daily share volume, daily dollar volume, or daily turnover in month $t-1$, $t-2$, or in the past 36 months, as well as monthly share volume, monthly dollar volume, or monthly turnover in the past 60 months. Then we have estimated the two illiquidity measures using a number of combinations of these as inputs. However, the cross-sectional regressions using liquidity 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.

5 Conclusion

Empirical proxies for illiquidity have been subject to controversy because they have achieved mixed results in the analysis of the issue of whether illiquidity is related to asset returns. Further, these proxies typically do not emanate from an equilibrium model, raising the question of whether the ad hoc empiricism that justifies the measures is the cause of conflicting conclusions about the illiquidity-return relation.

We use an alternative approach to measuring illiquidity. Specifically, we estimate Kyle lambdas, using analytic formulae that emanate from an economic 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 NYSE and Nasdaq stocks, spanning more than 35 years. We find convincing evidence that these theory-based lambdas are priced in the cross-section of expected stock returns.

Future research would focus on illiquidity-risk pricing and also why and how such theory-based illiquidity measures vary over time and across firms. For example, are such lambdas relatively high when the firm is young and there is a lot information asymmetry? Do they decline as the clientele for holding stocks changes from ostensibly informed institutions to uninformed individual investors? Can these lambdas be estimated for other markets such as fixed income, and do these vary with credit risk (presumably because the potential for asymmetric information is greater in bonds with high default potential)? Such issues are left for future research.

References

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

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

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

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

Avramov, D., T. Chordia, and A. Goyal, 2004, Liquidity and autocorrelations in individual stock returns, forthcoming, *Journal of Finance*.

Badrinath, S., J. Kale, and T. Noe, 1995, Of shepherds, sheep, and the cross-autocorrelations in equity returns, *Review of Financial Studies* 8, 401-430.

Ball, C., and T. Chordia, 2001, True spreads and equilibrium prices, *Journal of Finance* 56, 1801-1836.

Benston, G., and R. Hagerman, 1974, Determinants of bid-asked spreads in the over-the-counter market, *Journal of Financial Economics* 1, 353-364.

Bessembinder, H., 2003, Trade execution costs and market quality after decimalization, *Journal of Financial and Quantitative Analysis* 38, 747-777.

Borio, C., and R. McCauley, 1996, The economics of recent bond yield volatility, BIS Economic Paper No. 44, Bank for International Settlements, Basel, Switzerland.

Branch, B., and W. Freed, 1977, Bid-asked spreads on the AMEX and the big board, *Journal*

of Finance 32, 159-163.

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, 361-381.

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

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, 345-373.

Chan, K., 1993, Imperfect information and cross-autocorrelation among stock prices, *Journal of Finance* 48, 1211-1230.

Chan, K., and W. Fong, 2000, Trade size, order imbalance, and the volatility-volume relation, *Journal of Financial Economics* 57, 247-273.

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, 2001, Market liquidity and trading activity, *Journal of Finance* 56, 501-530.

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

Chordia, T., A. Sarkar, and A. Subrahmanyam, 2005, An empirical analysis of stock and bond market liquidity, *Review of Financial Studies* 18, 85-129.

Chordia, T., L. Shivakumar, and A. Subrahmanyam, 2004, Liquidity dynamics across small and large firms, *Economic Notes* 33, 111-143.

Chordia, T., and A. Subrahmanyam, 1995, Market making, the tick size, and payment-for-order flow: Theory and evidence, *Journal of Business* 68, 543-575.

Chordia, T., and B. Swaminathan, 2000, Trading volume and cross-autocorrelations in stock returns, *Journal of Finance* 55, 913-935.

Chowdhry, B., and V. Nanda, 1991, Multimarket trading and market liquidity, *Review of Financial Studies* 4, 483-511.

Chung K., and C. Chuwonganant, 2004, Tick size, order handling rules, and trading costs, *Financial Management* 33, 47-62.

Clark, R., J. McConnell, and M. Singh, 1992, Seasonalities in NYSE bid-ask spreads and stock returns in January, *Journal of Finance* 47, 1999-2014.

Conrad, J., M. Gultekin, and G. Kaul, 1991, Asymmetric predictability of conditional variances, *Review of Financial Studies* 4, 597-622.

Davis, P., 1999, A reappraisal of market liquidity risk in the light of the Russia/LTCM global securities market crisis, working paper, Bank of England, London, U.K.

Dennis, P., and D. Strickland, 2003, Who blinks in volatile markets? Individuals or Institutions? *Journal of Finance* 53, 1923-1949.

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

Emmons, W., and F. Schmid, 2000, The Asian crisis and the exposure of large U.S. firms, Federal Reserve Bank of St. Louis Review, January 2000, 15-34.

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

French, K., 1980, Stock returns and the weekend effect, *Journal of Financial Economics* 8, 55-69.

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

Gibbons, M., and R. Hess, 1981, Day of the week effects and asset returns, *Journal of Business* 54, 579-596.

Glosten, L., and P. Milgrom, 1985, Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71-100.

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

Gorton, G., and G. Pennacchi, 1993, Security baskets and index-linked securities, *Journal of Business* 66, 1-27.

Harris, L., 1995, Minimum price variations, discrete bid-ask spreads, and quotation sizes, *Review of Financial Studies* 7, 149-178.

Hasbrouck, J., 1991, Measuring the information content of stock trades, *Journal of Finance* 56, 179-207.

Hasbrouck, J., and D. Seppi, 2001, Common factors in prices, order flows and liquidity, *Journal of Financial Economics* 59, 383-411.

Haugen, R., and J. Lakonishok, 1987, *The Incredible January Effect*, Dow Jones-Irwin, Homewood, IL.

Hiemstra, C., and J. Jones, 1994, Testing for linear and nonlinear Granger causality in the stock price-volume relation, *Journal of Finance* 49, 1639-1664.

Hirshleifer, D., A. Subrahmanyam, and S. Titman, 1994, Security analysis and trading

patterns when some investors receive information before others, *Journal of Finance* 49, 1665-1698.

Ho, T., and H. Stoll, 1983, The dynamics of dealer markets under competition, *Journal of Finance* 38, 1053-1074.

Hou, K., and T. Moskowitz, 2005, Market frictions, price delay, and the cross-section of expected returns, *Review of Financial Studies* 18, 981-1020.

Huang, R., and H. Stoll, 1996, Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE, ıJournal of Financial Economics 41, 313-357.

Huberman, G., and D. Halka, 2001, Systematic liquidity, *Journal of Financial Research* 24, 161-178.

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.

Johnson, T., 2005, Dynamic liquidity in endowment economies, forthcoming, *Journal of Financial Economics.*

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

Jones, C., G. Kaul, and M. Lipson, 1994, Transactions, volume, and volatility, *Review of Financial Studies* 7, 631-651.

Keim, D., 1983, Size-related anomalies and stock return seasonality: Further empirical evidence, *Journal of Financial Economics* 12, 13-32.

Kumar, A., and C. Lee, 2005, Retail investor sentiment and return comovements, forthcoming, *Journal of Finance.*

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

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

Leach, C., and A. Madhavan, 1993, Price experimentation and security market structure, *Review of Financial Studies* 6, 375-404.

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

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

Lee, C., A. Shleifer, and R. Thaler, 1991, Investor sentiment and the closed-end fund puzzle, *Journal of Finance* 46, 75-109.

Lehmann, B., 1990, Fads, martingales, and market efficiency, *Quarterly Journal of Economics* 105, 1-28.

Lo, A., and C. MacKinlay, 1990, When are contrarian profits due to stock market overreaction?, *Review of Financial Studies* 3, 175-205.

Lo, A., and J. Wang, 2000, Trading volume: definitions, data analysis, and implications of portfolio theory, *Review of Financial Studies* 13, 257-300.

Longin, F., 1997, The threshold effect in expected volatility: A model based on asymmetric information, *Review of Financial Studies* 10, 837-869.

McQueen, G., M. Pinegar, and S. Thorley, 1996, Delayed reaction to good news and the cross-autocorrelation of portfolio returns, *Journal of Finance* 51, 889-919.

Mech, T., 1993, Portfolio return autocorrelation, *Journal of Financial Economics* 34, 307-344.

Merton, R., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867-887.

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

Mitchell, M., T. Pulvino, and E. Stafford, 2002, Limited arbitrage in equity markets, *Journal of Finance* 57, 551-584.

Muller, A., and W. Verschoor, 2004, The financial crisis exchange rate risk of U.S. multinationals, working paper, Limburg Institute of Financial Economics, Maastricht, Netherlands.

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

Odean, T., 1998, Are investors reluctant to realize their losses?, *Journal of Finance* 53, 1775-1798.

O'Hara, M., and G. Oldfield, 1986, The microeconomics of market making, *Journal of Financial and Quantitative Analysis* 21, 361-376.

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

Pesaran, H., and Y. Shin, 1998, Generalised impulse response analysis in linear multivariate models, *Economics Letters* 58, 17-29.

Rozeff, M., and W. Kinney, 1976, Capital market seasonality: The case of stock returns, *Journal of Financial Economics* 3, 379-402.

Sadka, R., and A. Scherbina, 2004, Analyst disagreement, mispricing, and liquidity, working paper, Northwestern University.

Schwert, W., 1990, Stock market volatility and the crash of 87, 1990, *Review of Financial*

Studies 3, 77-102.

Sias, R., and L. Starks, 1997, Return autocorrelation and institutional investors, *Journal of Financial Economics* 46, 103-131.

Stoll, H., 1978, The pricing of security dealer services: An empirical study of NASDAQ stocks, *Journal of Finance* 33, 1153-1172.

Stoll, H., 2000, Presidential address: Friction, *Journal of Finance* 55, 1479-1514.

Subrahmanyam, A., 1991, A theory of trading in stock index futures, *Review of Financial Studies* 4, 17-51.

Subrahmanyam, A., 1994, Circuit breakers and market volatility: A theoretical perspective, *Journal of Finance* 49,237-254.

Subrahmanyam, A., and S. Titman, 1999, The going public decision and the development of financial markets, *Journal of Finance*, 54, 1045-1082.

Appendix: Derivation of the Two Theory-based Illiquidity Measures from the Strategic Microstructure Model with Many Informed Traders

In this appendix, we derive the illiquidity measure (λ) used in our analysis, assuming that there are many informed traders in the Kyle's (1985) setting.

Consider an asset which 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 observes 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, ..., 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 noise

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 clear.

At each auction, 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. He 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

$$x_j = \gamma(\tilde{\delta} + \tilde{\varepsilon}_j), \tag{13}$$

and also that for informed trader i is

$$x_i = \gamma(\tilde{\delta} + \tilde{\varepsilon}_i). \tag{14}$$

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

$$\begin{aligned} \omega &= \{x_i + (N-1)x_j\} + z \\ &= \{x_i + (N-1)\gamma\tilde{\delta} + \gamma \sum_{j \neq i} \tilde{\varepsilon}_j\} + z \qquad (15) \\ &= N\gamma\tilde{\delta} + \gamma \sum_i \tilde{\varepsilon}_i + z. \qquad (16) \end{aligned}$$

The asset price, P, is set by the market maker after he observes ω so that[13]

$$P = E\left[\overline{W} + \tilde{\delta}|\omega = \mathrm{N}\gamma\tilde{\delta} + \gamma\sum_{\mathrm{i}}\tilde{\varepsilon}_{\mathrm{i}} + \mathrm{z}\right] \tag{17}$$

$$= \overline{W} + \frac{Cov\left(\tilde{\delta}, \mathrm{N}\gamma\tilde{\delta} + \gamma\sum_{\mathrm{i}}\tilde{\varepsilon}_{\mathrm{i}} + \mathrm{z}\right)}{Var\left(N\gamma\tilde{\delta} + \gamma\sum_{i}\tilde{\varepsilon}_{i} + z\right)}\left(N\gamma\tilde{\delta} + \gamma\sum_{i}\tilde{\varepsilon}_{i} + z\right). \tag{18}$$

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

$$P = \overline{W} + \lambda\omega \tag{19}$$

$$= \overline{W} + \lambda\left(N\gamma\tilde{\delta} + \gamma\sum_{i}\tilde{\varepsilon}_{i} + z\right), \tag{20}$$

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

$$\pi_i = (\tilde{W} - P)x_i. \tag{21}$$

In this setting, we can solve for the equilibrium that satisfies the 2 conditions: profit maximization by the informed, and market efficiency.

In Eq.(19) and Eq.(20), λ is an illiquidity measure and its inverse, $\frac{1}{\lambda}$, 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 out as a measure of illiquidity from the equilibrium conditions.

First, informed trader i's problem is:

$$Max \quad : \quad \mathrm{E}\left[\pi_{\mathrm{i}}|\tilde{\delta} + \tilde{\varepsilon}_{\mathrm{i}}\right]$$

[13]Equation (18) comes from the property of the multivariate normal distribution. Let 2 random variables, X_1 and X_2, be jointly normally distributed so that $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right]$. 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), for instance.

$$
\begin{aligned}
&= \mathrm{E}\left[\left\{\overline{\mathrm{W}}+\tilde{\delta}-\overline{\mathrm{W}}-\lambda\left(\mathrm{x_i}+(\mathrm{N}-1)\gamma\tilde{\delta}+\gamma\sum_{\mathrm{j}\neq\mathrm{i}}\tilde{\varepsilon}_{\mathrm{j}}\}+\mathrm{z}\right)\right\}\mathrm{x_i}\mid\tilde{\delta}+\tilde{\varepsilon}_{\mathrm{i}}\right] \\
&= \left(\mathrm{E}\left[\tilde{\delta}\mid\tilde{\delta}+\tilde{\varepsilon}_{\mathrm{i}}\right]-\lambda\mathrm{x_i}-\lambda(\mathrm{N}-1)\gamma\mathrm{E}\left[\tilde{\delta}\mid\tilde{\delta}+\tilde{\varepsilon}_{\mathrm{i}}\right]\right)\mathrm{x_i} \\
&= -\lambda x_i^2+x_i\left\{1-\lambda(N-1)\gamma\right\}E\left[\tilde{\delta}\mid\tilde{\delta}+\tilde{\varepsilon}_i\right]. \qquad (22)
\end{aligned}
$$

The first order condition of Eq.(22) gives $-2\lambda x_i+\{1-\lambda(N-1)\gamma\}E\left[\tilde{\delta}\mid\tilde{\delta}+\tilde{\varepsilon}_i\right]=0$. Thus,

$$
\begin{aligned}
x_i &= \frac{1}{2\lambda}\left\{1-\lambda(N-1)\gamma\right\}E\left[\tilde{\delta}\mid\tilde{\delta}+\tilde{\varepsilon}_i\right] \\
&= \frac{1}{2\lambda}\left\{1-\lambda(N-1)\gamma\right\}\frac{v_\delta}{v_\delta+v_\varepsilon}(\tilde{\delta}+\tilde{\varepsilon}_i). \qquad (23)
\end{aligned}
$$

Therefore, from Eq.(14) and Eq.(23), we have $\gamma=\frac{1}{2\lambda}\left\{1-\lambda(N-1)\gamma\right\}\frac{v_\delta}{v_\delta+v_\varepsilon}$, which in turn leads to

$$
\gamma=\frac{\frac{v_\delta}{v_\delta+v_\varepsilon}}{\lambda\left\{2+\frac{v_\delta}{v_\delta+v_\varepsilon}(N-1)\right\}}. \qquad (24)
$$

Next, from Eq.(18) and Eq.(20), the market efficiency condition is equivalent to

$$
\begin{aligned}
\lambda &= \frac{Cov\left(\tilde{\delta},\mathrm{N}\gamma\tilde{\delta}+\gamma\sum_{\mathrm{i}}\tilde{\varepsilon}_{\mathrm{i}}+\mathrm{z}\right)}{Var\left(N\gamma\tilde{\delta}+\gamma\sum_i\tilde{\varepsilon}_i+z\right)} \\
&= \frac{N\gamma v_\delta}{N^2\gamma^2v_\delta+\gamma^2Nv_\varepsilon+v_z} \\
&= \frac{Nv_\delta}{\gamma(N^2v_\delta+Nv_\varepsilon)+\frac{1}{\gamma}v_z}. \qquad (25)
\end{aligned}
$$

Plugging Eq.(24) into Eq.(25) gives

$$
\lambda=\frac{v_\delta}{(N+1)v_\delta+2v_\varepsilon}\sqrt{\frac{N(v_\delta+v_\varepsilon)}{v_z}}. \qquad (26)
$$

Note that initially we assumed informed traders observe a signal with a noise, $\tilde{\varepsilon}_i$ (i=1, 2, 3, ..., N). Now suppose there is no noise in the signal so that $v_\varepsilon=0$. Then, Eq.(26) is reduced to

$$
\lambda=\frac{\sqrt{Nv_\delta}}{(N+1)\sqrt{v_z}}. \qquad (27)
$$

In this study, Eq.(26) and Eq.(27) are used as the primary basis of our two illiquidity measures.

Table I

Descriptive Statistics for the Input Variables of the Two Illiquidity Measures

This table reports descriptive statistics for 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* = $P^{-1}\frac{v_\delta}{(N+1)v_\delta + 2v_\varepsilon}\sqrt{\frac{N(v_\delta + v_\varepsilon)}{v_z}}$, 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; v_δ: variance of payoff innovations; v_ε: variance of signal innovations; and v_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) 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 from four quarters ago; and *AVG(DVOL)-sqr*: squared value of AVG(DVOL). The sample periods are the past 372 months (31 years: 197201-200212) for NYSE/AMEX stocks and the 240 months (20 years: 198301-200212) for NASDAQ 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* in Panel A (NYSE/AMEX stocks) are 1,845.1 and 1,683.1, respectively. Those in Panel B (NASDAQ stocks) are 2,663.8 and 1,967.0, respectively.

Panel A: NYSE/AMEX

Original Input	Proxy Variables	For ILLIQ_1			For ILLIQ_2		
Variables	Employed	Mean	Median	STD	Mean	Median	STD
N	ANA	5.52	2.46	6.34	5.73	2.67	6.47
std(R)	STD(RET)	0.027	0.022	0.025	-	-	-
std(z)	AVG(DVOL)	5.26	0.53	16.88	-	-	-
P	P	-	-	-	31.31	18.06	379.92
v_δ	EVOLA-sqr	-	-	-	173.80	0.06	6298.46
v_ε	ESURP-sqr	-	-	-	69.68	0.02	2920.24
v_z	AVG(DVOL)-sqr	-	-	-	933.97	0.75	11061.09

Panel B: NASDAQ

Original Input	Proxy Variables	For ILLIQ_1			For ILLIQ_2		
Variables	Employed	Mean	Median	STD	Mean	Median	STD
N	ANA	3.15	1.37	4.01	3.70	1.79	4.56
std(R)	STD(RET)	0.044	0.036	0.039	-	-	-
std(z)	AVG(DVOL)	3.61	0.12	27.86	-	-	-
P	P	-	-	-	14.16	9.76	14.80
v_δ	EVOLA-sqr	-	-	-	1781.81	0.03	96989.02
v_ε	ESURP-sqr	-	-	-	3.22	0.01	73.02
v_z	AVG(DVOL)-sqr	-	-	-	2639.84	0.04	53269.12

Table II
Descriptive Statistics for Key Variables

This table reports descriptive statistics (Mean, Median, Standard Deviation (STD), Coefficient of Variation (CV), Skewness, and Kurtosis) for 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 I; *ILLIQ_2*: the second illiquidity measure defined as in Table I; *MV*: market value defined as the month-end stock price times the number of shares outstanding (in $million); *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); *BM_Trim*: the trimmed 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 3 months (from month t-1 to month t-3); *MOM2*: compounded holding period return over the next recent 3 months (from month t-4 to month t-6); *MOM3*: compounded holding period return over the 3 months from month t-7 to month t-9; *MOM4*: compounded holding period return over the 3 months from month t-9 to month t-12. The sample periods are the past 372 months (31 years: 197201-200212) for NYSE/AMEX stocks and the 240 months (20 years: 198301-200212) for NASDAQ 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 in Panel A (NYSE/AMEX stocks) is 1,845.1 (except that it is 1,683.1 for ILLIQ_2), while that in Panel B (NASDAQ stocks) is 2,663.8 (except that it is 1,967.0 for ILLIQ_2).

Panel A: NYSE/AMEX

Variables	Mean	Median	STD	CV	Skewness	Kurtosis
ILLIQ_1	2.14	0.11	14.92	686.22	17.29	441.55
ILLIQ_2	9.34	0.04	161.08	1229.57	24.13	730.84
MV	1722.92	245.01	6140.44	368.03	12.05	215.85
SIZE	5.15	5.13	1.98	39.10	0.09	-0.42
BM_Raw	1.02	0.87	0.78	78.08	4.92	80.07
BM_Trim	1.01	0.87	0.69	69.69	2.10	8.40
BTM	-0.27	-0.20	0.71	-387.46	-0.81	2.66
MOM1	0.011	0.016	0.191	-872.70	-0.27	5.55
MOM2	0.012	0.016	0.187	-1599.66	-0.18	4.94
MOM3	0.013	0.016	0.184	257.10	-0.10	4.41
MOM4	0.014	0.017	0.183	38.04	-0.06	4.14

Panel B: NASDAQ

Variables	Mean	Median	STD	CV	Skewness	Kurtosis
ILLIQ_1	10.25	0.18	271.22	1154.10	29.12	1130.67
ILLIQ_2	44.63	0.06	1272.14	1503.41	27.50	937.70
MV	323.01	47.56	2770.90	570.96	20.12	624.91
SIZE	3.84	3.76	1.65	43.49	0.24	-0.05
BM_Raw	0.75	0.58	0.71	93.08	4.23	49.95
BM_Trim	0.74	0.58	0.65	86.47	2.40	9.96
BTM	-0.66	-0.55	0.92	-147.06	-0.85	2.14
MOM1	-0.013	-0.001	0.289	-1496.29	-0.27	4.71
MOM2	-0.008	0.001	0.284	-11200.06	-0.16	4.05
MOM3	-0.004	0.004	0.280	64.58	-0.09	3.83
MOM4	-0.003	0.004	0.278	745.88	-0.06	3.78

Table III
Correlations between Explanatory Variables

The lower triangle shows the average correlations between the key variables for NYSE/AMEX stocks over the 372 months (31 years: 197201-200212), and the upper triangle shows those for NASDAQ stocks over the 240 months (20 years: 198301-200212) The 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 I; *ILLIQ_2*: the second illiquidity measure defined as in Table I; *SIZE*: natural logarithm of MV, which is defined as the month-end stock price times the number of shares outstanding (in $million); *BTM*: natural logarithm of BM_Trim, which is the trimmed 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 3 months (from month t-1 to month t-3); *MOM2*: compounded holding period return over the next recent 3 months (from month t-4 to month t-6); *MOM3*: compounded holding period return over the 3 months from month t-7 to month t-9; *MOM4*: compounded holding period return over the 3 months from month t-9 to month t-12. The average number of component stocks used in a month for NYSE/AMEX stocks is 1,845.1 (except that it is 1,683.1 for ILLIQ_2), while that for NASDAQ stocks is 2,663.8 (except that it is 1,967.0 for ILLIQ_2). Under the null hypothesis of zero correlation, asymptotic standard error of the correlation coefficient is $1/\sqrt{N}$, where N is the number of component stocks used in computing the correlation coefficients for each month.

Average Correlations: NYSE/AMEX (Lower Triangle) and NASDAQ (Upper Triangle)

	ILLIQ_1	ILLIQ_2	SIZE	BTM	MOM1	MOM2	MOM3	MOM4
ILLIQ_1	1	0.464	-0.182	0.072	-0.055	-0.049	-0.040	-0.036
ILLIQ_2	0.639	1	-0.154	0.065	-0.066	-0.048	-0.046	-0.036
SIZE	-0.302	-0.184	1	-0.247	0.176	0.173	0.172	0.171
BTM	0.091	0.064	-0.290	1	0.089	0.081	0.045	-0.009
MOM1	-0.088	-0.074	0.122	0.043	1	0.019	0.045	0.037
MOM2	-0.079	-0.059	0.120	0.039	0.015	1	0.014	0.040
MOM3	-0.072	-0.056	0.121	0.009	0.047	0.012	1	0.007
MON4	-0.073	-0.054	0.122	-0.034	0.052	0.044	0.006	1

Table IV

Average Values of Monthly Return, Size, and Illiquidity for the 25 Portfolios Formed on Illiquidity and Size

This table reports the average values of monthly stock return (Panel A), firm size (Panel B), and illiquidity (Panel C) for the 25 portfolios formed on illiquidity and size. ILLIQ_1 is the illiquidity measure defined as in Table I and size is the market value (MV) of a firm (in $million). The component stocks are first split into 5 portfolios (with the equal number of stocks) after being sorted in an ascending order by ILLIQ_1 and then each of the 5 portfolios is again split into another 5 portfolios after being sorted by size, resulting in 25 portfolios each month. The average values of return, size, and illiquidity are computed each month for each of the 25 portfolios, and then the time-series averages of the 3 variables over the sample period are reported in each panel of the table. Panel A also contains t-statistics (italicized), in addition to the average returns. The sample periods are the past 372 months (31 years: 197201-200212) for NYSE/AMEX stocks and the 240 months (20 years: 198301-200212) for NASDAQ stocks. The average number of component stocks in each portfolio in a month for NYSE/AMEX stocks is 73.4, while that for NASDAQ stocks is 106.1.

Panel A: Average Return

	Size Group (NYSE/AMEX)					Size Group (NASDAQ)				
ILLIQ_1 Group	1 small	2	3	4	5 big	1 small	2	3	4	5 big
1 low	0.0118	0.0116	0.0104	0.0096	0.0093	0.0110	0.0100	0.0089	0.0083	0.0081
	3.94	*4.29*	*4.07*	*3.95*	*3.91*	*2.85*	*2.27*	*2.28*	*2.19*	*2.36*
2	0.0120	0.0124	0.0118	0.0103	0.0103	0.0110	0.0101	0.0076	0.0080	0.0079
	3.45	*4.45*	*4.09*	*3.72*	*4.21*	*1.90*	*2.17*	*2.02*	*2.29*	*3.22*
3	0.0120	0.0125	0.0119	0.0110	0.0103	0.0131	0.0101	0.0078	0.0089	0.0088
	3.35	*3.72*	*3.31*	*4.03*	*3.80*	*2.28*	*2.10*	*1.61*	*2.16*	*2.61*
4	0.0136	0.0127	0.0124	0.0112	0.0104	0.0207	0.0107	0.0078	0.0091	0.0095
	3.54	*3.54*	*3.79*	*3.56*	*3.66*	*3.31*	*2.27*	*1.80*	*1.79*	*2.37*
5 high	0.0291	0.0127	0.0125	0.0114	0.0101	0.0640	0.0244	0.0141	0.0090	0.0089
	5.27	*2.80*	*2.55*	*2.40*	*2.47*	*8.70*	*4.58*	*3.03*	*2.02*	*1.27*

Panel B: Average Size ($million)

	Size Group (NYSE/AMEX)					Size Group (NASDAQ)				
ILLIQ_1 Group	1 small	2	3	4	5 big	1 small	2	3	4	5 big
1 low	1194.29	2179.59	3397.65	5898.20	23237.69	126.85	270.57	437.02	769.11	5160.62
2	317.96	528.95	739.70	1034.03	2008.78	46.09	79.50	113.26	164.76	335.73
3	104.28	172.21	240.05	338.44	656.28	18.69	32.40	46.13	66.94	150.88
4	30.00	52.66	77.31	112.35	238.31	7.93	14.21	20.77	30.49	69.87
5 high	4.31	9.46	15.65	24.84	61.47	2.36	4.80	7.56	11.97	29.73

Panel C: Average ILLIQ_1

	Size Group (NYSE/AMEX)					Size Group (NASDAQ)				
ILLIQ_1 Group	1 small	2	3	4	5 big	1 small	2	3	4	5 big
1 low	0.008	0.007	0.005	0.003	0.001	0.006	0.008	0.006	0.004	0.002
2	0.037	0.033	0.029	0.024	0.019	0.056	0.047	0.041	0.037	0.033
3	0.145	0.126	0.114	0.102	0.091	0.245	0.216	0.196	0.181	0.166
4	0.683	0.574	0.485	0.418	0.357	1.106	0.965	0.891	0.826	0.752
5 high	31.766	8.288	4.703	3.175	2.436	173.772	32.932	15.874	20.877	7.853

Table V

Results of Monthly Cross-sectional Regressions: with ILLIQ_1 for NYSE/AMEX and NASDAQ Stocks

This table reports the monthly Fama-MacBeth (1973)-type cross-sectional regressions using our first measure of illiquidity, ILLIQ_1, for NYSE/AMEX stocks over the 372 months (31 years: 197201-200212) and for NASDAQ stocks over the 240 months (20 years: 198301-200212). The dependent variables (EXSRET, FF3-adj EXSRET1, and FF3-adj EXSRET2) are all one-month leading values (no contemporaneous regressors are used). 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) 3 factors, i.e., the constant term plus the residual from the time-series regression of the excess return on the FF 3 factors using the *whole* sample range of the data; *FF3-adj EXSRET2*: the risk-adjusted excess return using the Fama-French (FF) 3 factors, i.e., R_i^* computed each month with the current month data from the equation, $R_i^* = (R_i - R_f) - [\hat{\beta}_1 MKT + \hat{\beta}_2 SMB + \hat{\beta}_3 HML]$, after the factor loadings ($\alpha, \beta_1, \beta_2, \beta_3$) are first estimated for *each month* using the time-series data of the past 60 months in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \varepsilon$, where R_i, R_f, and R_m are the individual stock return, the risk-free rate, and the market index return, respectively, while *MKT*, *SMB*, and HML are FF 3 factors; *ILLIQ_1*: the first illiquidity measure 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 I); *SIZE*: natural logarithm of MV, which is defined as the month-end stock price times the number of shares outstanding (in $million); *BTM*: natural logarithm of BM_Trim, which is the trimmed 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 3 months (from month t-1 to month t-3); *MOM2*: compounded holding period return over the next recent 3 months (from month t-4 to month t-6); *MOM3*: compounded holding period return over the 3 months from month t-7 to month t-9; *MOM4*: compounded holding period return over the 3 months from month t-9 to month t-12. The average number of component stocks used in the monthly regressions for NYSE/AMEX stocks is 1,845.1, while that for NASDAQ stocks is 2,663.8. The values in the first row for each explanatory variable under the columns labeled '*Unadjusted*' and '*Raw*' 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 values under the columns labeled '*Purged*' provide the purged estimates based on Black et al. (1972), which are the constant terms and their t-values from the regressions of the monthly coefficient estimates on the time-series of the FF 3 factor realizations. 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.

(Table V continued)

	NYSE/AMEX										NASDAQ									
Explanatory	EXSRET		FF3-adj EXSRET1				FF3-adj EXSRET2				EXSRET		FF3-adj EXSRET1				FF3-adj EXSRET2			
Variables	Unadjusted		Raw		Purged		Raw		Purged		Unadjusted		Raw		Purged		Raw		Purged	
Intercept	0.994	**	0.187		0.119		0.239		0.154		2.337	**	1.694	**	1.633	**	1.759	**	1.677	**
	2.65		*1.19*		*0.78*		*1.40*		*0.92*		*4.75*		*4.58*		*4.86*		*4.50*		4.72	
ILLIQ_1	0.044	**	0.039	**	0.039	**	0.041	**	0.041	**	0.035	**	0.033	**	0.033	**	0.033	**	0.033	**
	3.55		*3.18*		*3.24*		*3.12*		*3.18*		*3.94*		*3.97*		*3.83*		*3.87*		3.74	
SIZE	-0.096	*	-0.053	*	-0.047	*	-0.053	*	-0.047		-0.351	**	-0.325	**	-0.317	**	-0.325	**	-0.315	**
	-2.15		*-2.19*		*-1.99*		*-2.08*		*-1.85*		*-4.25*		*-4.43*		*-4.64*		*-4.17*		-4.35	
BTM	0.223	**	0.161	**	0.139	**	0.102	*	0.087		0.323	**	0.341	**	0.303	**	0.308	**	0.282	**
	3.17		*3.40*		*2.88*		*2.27*		*1.90*		*2.87*		*4.34*		*3.94*		*3.37*		3.11	
MOM1	0.863	**	0.938	**	0.925	**	0.907	**	0.888	**	0.324		0.431		0.459		0.333		0.332	
	2.71		*3.63*		*3.55*		*2.94*		*2.85*		*0.93*		*1.66*		*1.72*		*0.87*		0.85	
MOM2	1.177	**	1.144	**	1.184	**	1.043	**	1.078	**	0.677	*	0.495	*	0.550	*	0.492		0.547	
	4.31		*5.12*		*5.20*		*3.91*		*3.96*		*2.35*		*2.18*		*2.39*		*1.61*		1.76	
MOM3	1.435	**	1.373	**	1.372	**	1.112	**	1.159	**	0.613	*	0.473	*	0.557	**	0.294		0.478	
	5.47		*6.04*		*5.92*		*4.10*		*4.21*		*2.54*		*2.20*		*2.58*		*0.95*		1.55	
MOM4	1.006	**	0.961	**	1.036	**	0.720	**	0.841	**	0.394		0.343		0.418	*	0.209		0.315	
	4.76		*4.97*		*5.29*		*2.77*		*3.21*		*1.83*		*1.90*		*2.29*		*0.75*		1.13	
Avg R-sqr	0.049		0.029				0.032				0.034		0.023				0.028			
Avg Obs	1845.1		1843.4				1797.7				2663.8		2658.9				2406.3			

Table VI

Results of Monthly Cross-sectional Regressions: with ILLIQ_2 for NYSE/AMEX and NASDAQ Stocks

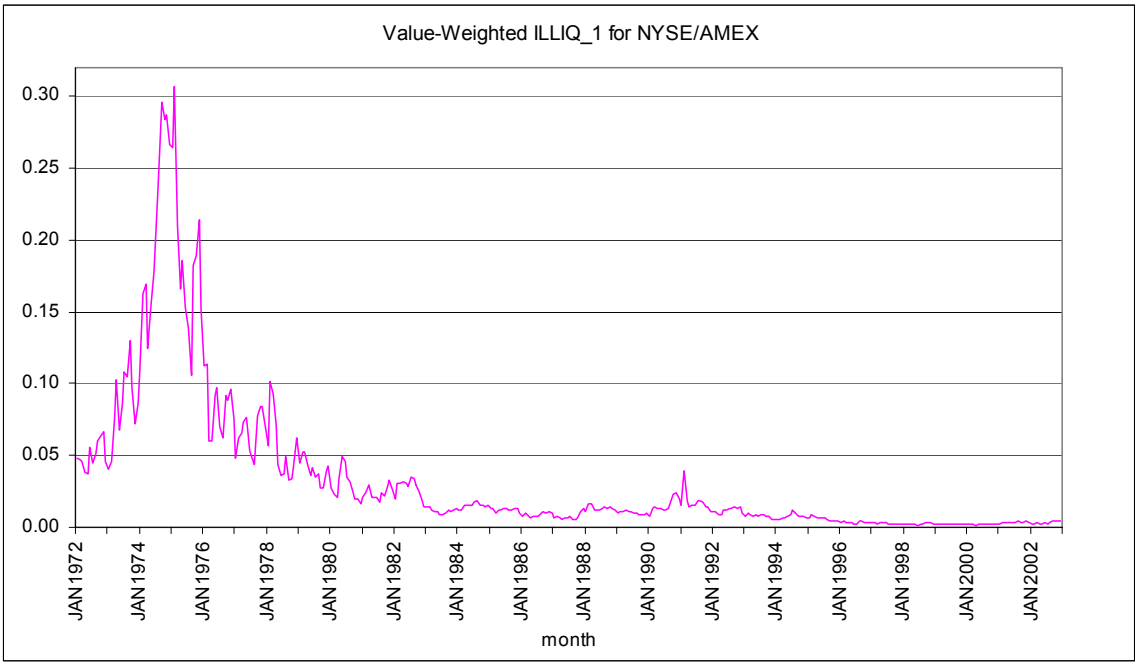
This table reports the monthly Fama-MacBeth (1973)-type cross-sectional regressions using our second measure of illiquidity, ILLIQ_2, for NYSE/AMEX stocks over the 372 months (31 years: 197201-200212) and for NASDAQ stocks over the 240 months (20 years: 198301-200212). The dependent variables (EXSRET, FF3-adj EXSRET1, and FF3-adj EXSRET2) are all one-month leading values (no contemporaneous regressors are used). 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) 3 factors, i.e., the constant term plus the residual from the time-series regression of the excess return on the FF 3 factors using the *whole* sample range of the data; *FF3-adj EXSRET2*: the risk-adjusted excess return using the Fama-French (FF) 3 factors, i.e., R_i^* computed each month with the current month data from the equation, $R_i^* = (R_i - R_f) - [\hat{\beta}_1 MKT + \hat{\beta}_2 SMB + \hat{\beta}_3 HML]$, after the factor loadings ($\alpha, \beta_1, \beta_2, \beta_3$) are first estimated for *each month* using the time-series data of the past 60 months in the monthly regression, $R_i - R_f = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \varepsilon$, where R_i, R_f, and R_m are the individual stock return, the risk-free rate, and the market index return, respectively, while *MKT*, *SMB*, and HML are FF 3 factors; *ILLIQ_2*: the second illiquidity measure defined as $P^{-1} \frac{v_\delta}{(N+1)v_\delta + 2v_\varepsilon} \sqrt{\frac{N(v_\delta + v_\varepsilon)}{v_z}}$, where *P* is the asset price, *N* is the number of informed traders, v_δ is the variance of payoff innovations, v_ε is the variance of signal innovations, and v_z is the variance of noise trades (the original input variables are proxied by the variables as shown in Table I); *SIZE*: natural logarithm of MV, which is defined as the month-end stock price times the number of shares outstanding (in $million); *BTM*: natural logarithm of BM_Trim, which is the trimmed 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 3 months (from month t-1 to month t-3); *MOM2*: compounded holding period return over the next recent 3 months (from month t-4 to month t-6); *MOM3*: compounded holding period return over the 3 months from month t-7 to month t-9; *MOM4*: compounded holding period return over the 3 months from month t-9 to month t-12. The average number of component stocks used in the monthly regressions for NYSE/AMEX stocks is 1,683.1, while that for NASDAQ stocks is 1,967.0. The values in the first row for each explanatory variable under the columns labeled '*Unadjusted*' and '*Raw*' 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 values under the columns labeled '*Purged*' provide the purged estimates based on Black et al. (1972), which are the constant terms and their t-values from the regressions of the monthly coefficient estimates on the time-series of the FF 3 factor realizations. 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.

(Table VI continued)

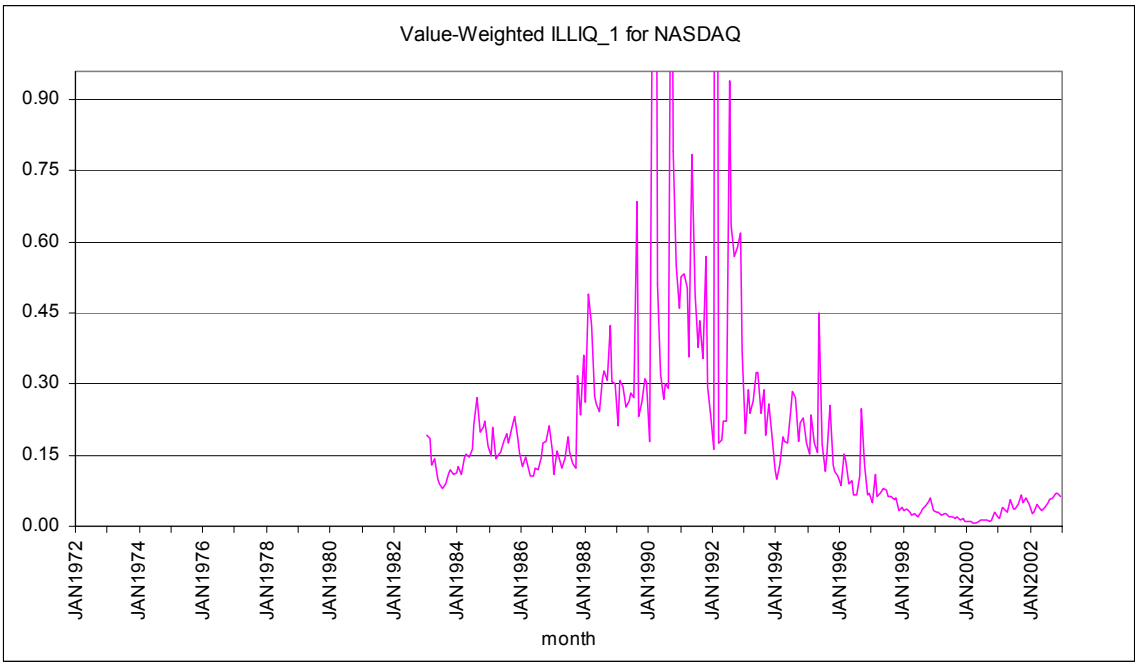
	NYSE/AMEX										NASDAQ									
Explanatory	EXSRET		FF3-adj EXSRET1				FF3-adj EXSRET2				EXSRET		FF3-adj EXSRET1				FF3-adj EXSRET2			
Variables	Unadjusted		Raw		Purged		Raw		Purged		Unadjusted		Raw		Purged		Raw		Purged	
Intercept	1.120	**	0.295		0.212		0.356	*	0.256		1.937	**	1.257	**	1.218	**	1.318	**	1.255	**
	2.91		*1.74*		*1.29*		*1.97*		*1.47*		*3.97*		*3.54*		*3.72*		*3.47*		*3.58*	
ILLIQ_2	0.010	**	0.008	*	0.009	**	0.009	*	0.010	**	0.008		0.004		0.003		0.004		0.003	
	2.84		*2.37*		*2.73*		*2.45*		*2.85*		*1.04*		*0.55*		*0.39*		*0.54*		*0.40*	
SIZE	-0.112	*	-0.064	*	-0.056	*	-0.068	*	-0.06	*	-0.242	**	-0.220	**	-0.217	**	-0.223	**	-0.218	**
	-2.42		*-2.44*		*-2.18*		*-2.46*		*-2.19*		*-3.09*		*-3.20*		*-3.36*		*-2.98*		*-3.10*	
BTM	0.184	**	0.134	**	0.116	*	0.073		0.061		0.366	**	0.361	**	0.317	**	0.349	**	0.314	**
	2.59		*2.74*		*2.34*		*1.58*		*1.29*		*2.98*		*4.15*		*3.69*		*3.60*		*3.25*	
MOM1	0.806	*	0.847	**	0.858	**	0.874	**	0.878	**	0.812	*	0.876	**	0.890	**	0.817	*	0.800	*
	2.51		*3.26*		*3.26*		*2.87*		*2.84*		*2.08*		*2.91*		*2.89*		*2.05*		*1.96*	
MOM2	1.068	**	1.054	**	1.112	**	1.004	**	1.057	**	0.960	**	0.795	**	0.849	**	0.707	*	0.771	*
	3.90		*4.70*		*4.88*		*3.82*		*3.94*		*3.05*		*2.99*		*3.14*		*2.14*		*2.28*	
MOM3	1.362	**	1.301	**	1.316	**	1.071	**	1.133	**	0.699	**	0.612	**	0.688	**	0.323		0.511	
	5.15		*5.66*		*5.61*		*3.91*		*4.07*		*2.73*		*2.58*		*2.90*		*1.04*		*1.66*	
MOM4	0.954	**	0.930	**	0.997	**	0.652	*	0.762	**	0.568	*	0.575	**	0.642	**	0.440		0.551	
	4.33		*4.57*		*4.83*		*2.43*		*2.81*		*2.20*		*2.59*		*2.83*		*1.47*		*1.82*	
Avg R-sqr	0.050		0.029				0.032				0.040		0.028				0.032			
Avg Obs	1683.1		1682.0				1667.6				1967.0		1961.6				1907.3			

Figure 1. Trends of the Value-weighted Illiquidity Measure: ILLIQ_1

The following graphs show the trends of the market value-weighted ILLIQ_1 measure over the past 372 months (31 years: 197201-200212). The value-weighted illiquidity series are the monthly cross-sectional (market-value weighted) averages of ILLQ_1 over the sample period. ILLIQ_1 is defined as in Table I. Figure 2(a) is for the stocks on the NYSE/AMEX, and Figure 2(b) for those on the NASDAQ (available from January 1983 to December 2002 only). The average numbers of component stocks used each month are 1,845.1 for NYSE/AMEX (197201-200212) stocks and 2,663.8 for NASDAQ (198301-200212) stocks.



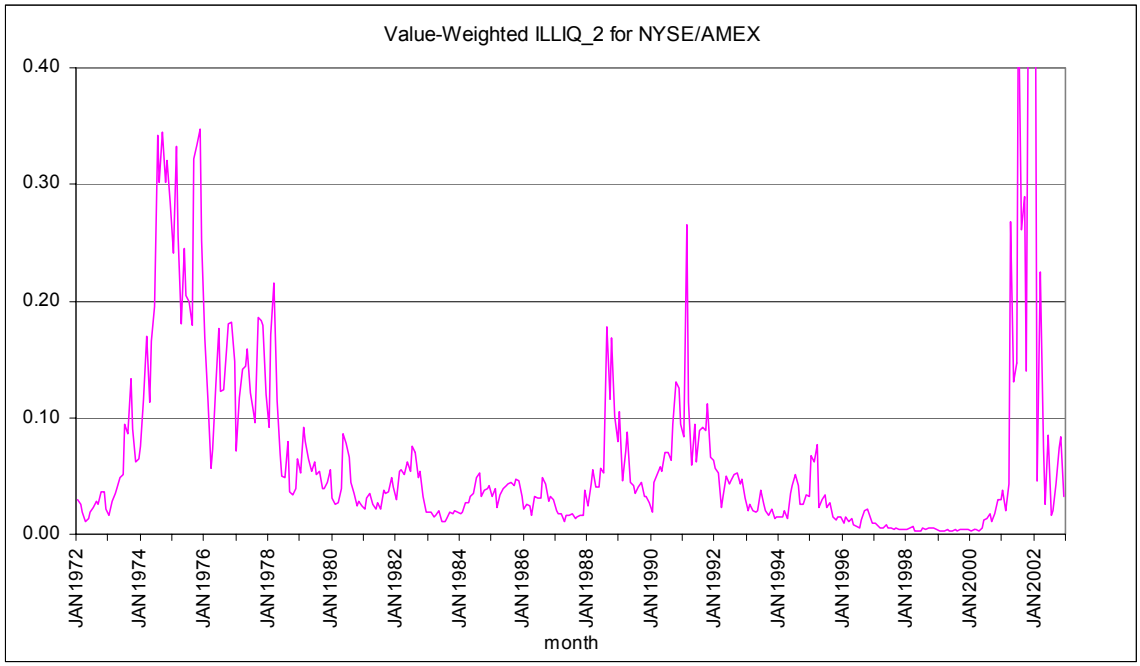
(a)



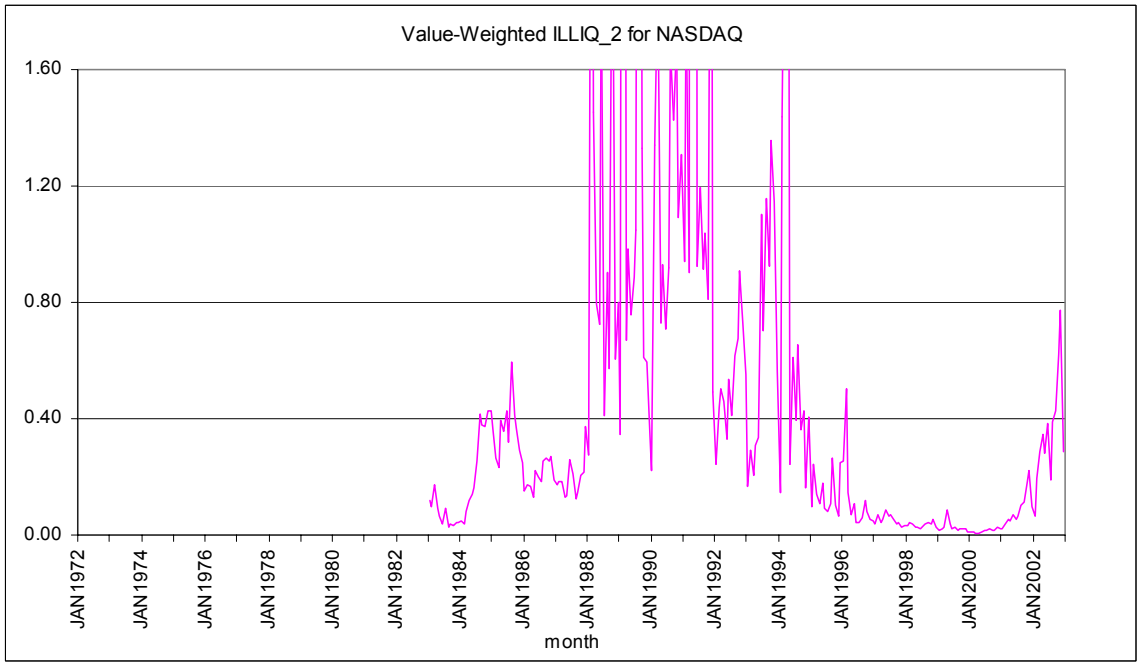
(b)

Figure 2. Trends of the Value-weighted Illiquidity Measure: ILLIQ_2

The following graphs show the trends of the market value-weighted ILLIQ_2 measure over the past 372 months (31 years: 197201-200212). The value-weighted illiquidity series are the monthly cross-sectional (market-value weighted) averages of ILLQ_1 over the sample period. ILLIQ_2 is defined as in Table I. Figure 2(a) is for the stocks on the NYSE/AMEX, and Figure 2(b) for those on the NASDAQ (available from January 1983 to December 2002 only). The average numbers of component stocks used each month are 1,683.1 for NYSE/AMEX (197201-200212) stocks and 1,967.0 for NASDAQ (198301-200212) stocks.



(a)



(b)