ABSTRACT
This paper explains and extends my 2002 paper. It presents a return factor of illiquid-minus-liquid stocks (IML), which provides a time-series of the illiquidity premium. The risk-adjusted expected return on IML is positive and significant in the last 63 years and while it is lower in the period that follows my 2002 paper it remains positive and significant. IML also has the predicted response to market illiquidity shocks. In particular, the relation between illiquidity shocks and stock returns is more negative for illiquid stocks even after my study period.

Keywords: Liquidity and asset pricing, Illiquidity measure, Liquidity shocks, Liquidity premium

JEL Codes: G10, G12

1 Introduction

I am honored that the Critical Finance Review has commissioned two teams of excellent scholars, Drienko et al. (2018) and Harris and Amato (2018), to replicate and extend my 2002 study. I thank the authors of these studies for agreeing to undertake this task and for their analysis.

I begin by explaining why I developed the illiquidity measure (ILLIQ), the average ratio of absolute returns to the dollar trading volume, a low-frequency measure that can be calculated from daily data on CRSP. It was developed out of necessity. The purpose of my 2002 article is indicated by its title, “Cross-section and time-series analysis.” I intended to study the time-series effects of illiquidity shocks on realized stock returns, a subject that has not been hitherto studied. Such a study necessitated a reasonably long time-series of a measure of illiquidity. At that time, fine illiquidity measures such as the effective bid-ask spread and Kyle’s (1985) λ could be calculated from intraday data for 13 years whereas ILLIQ
could be calculated for 34 years. I then used ILLIQ which I showed to be positively related to Kyle’s (1985) \( \lambda \)—a measure of price impact cost—and to the fixed cost of trading, both estimated from intraday data.\(^1\)

The results in the first part of my 2002 paper on the positive cross-section effect of ILLIQ on expected returns are not new. They repeat the results in Amihud and Mendelson (1986) and others.\(^2\) This part was needed to validate the use of average market ILLIQ for the time-series analysis in the second part of the paper which presented new results:

1. Expected market ILLIQ positively affect expected returns and shocks to market ILLIQ negatively affect the time-series of realized returns. This follows from two pieces of evidence: ILLIQ has positive effect on expected return, as documented in the first part of the paper, and the aggregate market ILLIQ is highly autoregressive which means that an increase in market ILLIQ is expected to persist for a while. This raises subsequent expected returns and as a result stock prices fall for given cash flows.\(^3\) This analysis follows that of the effect of market volatility on stock returns; see Merton (1980) and French et al. (1987).

2. The negative effect of illiquidity shocks on realized return is greater (more negative) for smaller, less liquid stocks. I propose (p. 53): “the greater sensitivity of small stocks to illiquidity means that these stocks are subject to greater illiquidity risk which, if priced, should result in higher illiquidity risk premium.” Illiquidity risk—the covariance of returns with illiquidity shocks—is found to be positively priced by Pastor and Stambaugh (2003) and Acharya and Pedersen (2005).

Another insight is that the effects of illiquidity on the market excess return “suggest that the stock excess return, usually referred to as “risk premium,” is in part a premium for stock illiquidity. This contributes to the explanation of the puzzle that the equity premium is too high. “The results mean that stock excess returns reflect not only the higher risk but also the lower liquidity of stocks compared to Treasury securities.” (p. 53)

\(^1\)I regress ILLIQ on \( \lambda \) and \( \psi \), the latter being the fixed-cost component related to the bid–ask spread, estimated by the method of Glosten and Harris (1988) and available for 1984 (Michael Brennan and Avanidhar Subrahmanya kindly provided me with their estimates of these parameters). The coefficients of \( \lambda \) and \( \psi \) are positive and highly significant. Lesmond (2005) finds that ILLIQ is among the best proxies of within-country illiquidity measured by bid-ask spread plus commissions. Goyenko et al. (2009) and Hasbrouck (2009) show that ILLIQ performs best among low-frequency measures of \( \lambda \). Fong et al. (2017) find that ILLIQ is among the best low-frequency measures of \( \lambda \) across 39 global markets.

\(^2\)See Brennan and Subrahmanya (1996), Brennan et al. (1998), and Datar et al. (1998) and other studies, reviewed in Amihud et al. (2005, 2013).

\(^3\)The assumption is that illiquidity shocks do not affect cash flow. This condition can be weaker, see Acharya and Pedersen (2005).
There is earlier evidence on the time-series effect of illiquidity changes on stock prices in Amihud et al. (1990, 1997, 1999). Amihud et al. (1990) find that during the October 1987 stock market crash stock price declines were greater for stocks whose illiquidity increased by more. The issue of causality is resolved in the next two studies which show that exogenous declines in illiquidity cause price increases. Amihud et al. (1997) study the effects of a change in the market trading method in Israel where stocks selected by the exchange were transferred from an illiquid once-a-day auction to a semi-continuous trading system. The liquidity of the transferred stocks increased and so did their price. Amihud et al. (1999) study the effects of an increase in the investor base in the Tokyo Stock Exchange, achieved by a reduction in the minimum trading unit which facilitated trading by individual investors. This change induced an increase in liquidity of the affected stocks and an increase in their price.

ILLIQ is obviously not the only possible low-frequency measure of (il)liquidity. Another common liquidity measure is trading volume, a component of ILLIQ, or turnover. Evidence shows that the effect of volume or turnover on expected return in cross-section tests is negative and significant; see Brennan et al. (1998), Datar et al. (1998), and Chordia et al. (2001). Turnover is also included in the analysis in Amihud (2000) and there too its effect on expected return is negative and significant in addition to the positive and significant effect of ILLIQ. However, the time-series of aggregate market volume sometimes gives an incorrect reading of market liquidity. During the October 19, 1987 stock market crash, while illiquidity sharply increased as documented in the Brady report and in Amihud et al. (1990), trading volume rose too. And, during the recent financial crisis in September to October of 2008 market liquidity worsened, ILLIQ rose sharply as did other measures of illiquidity (bid–ask spread and Kyle’s λ), yet aggregate market volume remained flat. In Amihud (2000) I replicate the time-series tests using the series of market turnover (the cross-section average of stock turnover) instead of ILLIQ and find that while the results are qualitatively similar to those using market ILLIQ, when including both market ILLIQ and market turnover in the time-series model, “only illiquidity is statistically significant whereas turnover is not” (p. 28).

My objective in the 2002 article was to have a single measure of illiquidity that provides consistent results and performs well in both cross-section and time-series analyses and that could be (easily) obtained from widely available databases for

---


5 On October 19, 1987, the trading volume of Standard & Poor's (S&P) 500 stocks was 604.3 million shares, compared with the 141.9 million shares that traded a week before, on October 12, 1987. This is more than a fourfold increase. The S&P 500 index levels were 224.84 and 309.39 on October 19 and 12, 1987, respectively, suggesting that the dollar volume increased as well on October 19. Source: finance.yahoo.com.

6 Amihud and Noh (2018) discuss the inconsistency between the behavior of aggregate market volume and observed market liquidity. The pattern of a rise in market volume when illiquidity rises is noted by Pastor and Stambaugh (2003).
long periods of time. \textit{ILLIQ} satisfied this requirement, though other measures of illiquidity are not excluded and could perform as well. The key question is not which illiquidity measure to use but whether illiquidity—however measured—affects stock returns in a way consistent with the theory. As stated in my 2002 study, there may be finer measures of illiquidity. It is hard to expect that a single measure would include all aspects of liquidity. What is important, in my view, is the question of whether illiquidity is priced, regardless of the specific measure—or a combination of measures—that is used.

In what follows, I use \textit{ILLIQ} to present new analyses on the cross-sectional and time-series effects of illiquidity on stock returns.

\section{Construction of the Illiquid-Minus-Liquid Factor}

I present evidence on the illiquidity premium across stocks, using a return factor denoted illiquid-minus-liquid (ILML), the differential return on illiquid-minus-liquid stock portfolios.\textsuperscript{7} The illiquidity of stock $j$ on day $d$ is measured by $\text{ILLIQ}_{j,d} = |\text{return}_{j,d}|/\text{dollar volume}_{j,d}$ and is averaged over a 12-month period that ends in November of each year $y$. The variable $\text{ILLIQ}_{j,y}$ is used to analyze stock returns in year $y + 1$, as in Amihud (2002). In calculating annual $\text{ILLIQ}_{j,y}$ values, I delete stock–days with a negative price,\textsuperscript{8} a trading volume of less than 100 shares, or a return of less than $-100\%$ and I delete the highest daily value of $\text{ILLIQ}_{j,d}$ in each year. A stock is included if, during the 12-month period, its price is between $5$ and $1,000$ and it has more than 200 days of valid return and volume data. Finally, the sample in each year $y$ excludes stocks whose $\text{ILLIQ}_{j,y}$ values are in the top 1\%, since they are potential outliers. In addition, $\text{SD}_{j,y}$ is the standard deviation of the daily returns of stock $j$ over the same 12 months. I employ New York Stock Exchange (NYSE)/American Stock Exchange (AMEX) common stocks (codes 10 and 11).

Portfolios are formed in each month $t$ (January to December) in year $y$ for stocks that satisfy the above criteria and exist at the end of the preceding month. Stocks are sorted on $\text{SD}_{j,y-1}$ into three portfolios and, within each volatility portfolio, they are sorted by $\text{ILLIQ}_{j,y-1}$ into five portfolios, resulting in 15 ($3 \times 5$) portfolios. I do the double sorting\textsuperscript{9} because these two variables are positively correlated (Stoll, 1978; Amihud, 2002), each having its own effect on expected returns. In my 2002 cross-section analysis I control for volatility by including $\text{SD}_{j,y-1}$ among the explanatory variables and find that its effect on expected returns is

\textsuperscript{7}This factor is used in Amihud et al. (2013, 2015) and Amihud and Noh (2017).

\textsuperscript{8}This indicates that the price is the mid-point between the quoted bid and ask prices rather than a transaction price.

\textsuperscript{9}This procedure follows that of Fama and French (1993), where \textit{HML} is constructed by double-sorting stocks into size and book-to-market portfolios so as not to confound the effects of these two stock characteristics.
negative, as also found in Ang et al. (2006, 2009). I then calculate the monthly weighted average return for each portfolio using the capitalizations of the previous month as weights. Stock returns are adjusted by Shumway’s (1997) method to correct for the delisting bias. Finally, $IML_t$ is the average of the returns of month $t$ of the highest $ILLIQ$ quintile portfolios across the three corresponding SD portfolios minus the average returns on the lowest $ILLIQ$ quintile portfolios across the three corresponding SD portfolios.

### 3 The Risk-Adjusted Premium on $IML$

Table 1 presents estimated statistics of $IML$ for two periods: Period I, 1964 to 1997 (408 months) is the period studied in my 2002 article and Period II, 1998 to 2017 (240 months), extends the analysis until the present.

Panel A of Table 1 presents the mean, median, and proportion of months with positive values of $IML$. The mean $IML$ value is positive and significant being 0.635 ($t = 4.47$) and 0.430 ($t = 2.14$) in Periods I and II, respectively. The respective medians are 0.615 and 0.218, indicating positive skewness in Period II. In both periods, the proportion of months with $IML > 0$ is significantly greater than 0.50, the chance result.

Panel B of Table 1 presents $alpha$, the risk-adjusted mean $IML$, which is the intercept from a regression of $IML_t$ on the risk factors of Fama and French (1993) and Carhart (1997) (FFC):

$$IML_t = alpha + \beta_{RMrf} * RMrf_t + \beta_{SMB} * SMB_t$$

$$+ \beta_{HML} * HML_t + \beta_{UMD} * UMD_t + \epsilon_t$$ (1)

$RMrf$, $SMB$, $HML$, and $UMD$ are, respectively, the market excess return over the T-bill rate and the returns on small-minus-big firms (size factor), high-minus-low book-to-market (BE/ME) ratio firms (value-growth factor), and winner-minus-loser stocks (momentum factor). Panel B1 presents the results using only the factor $RMrf$, because firm size, used to construct $SMB$, is considered a measure of liquidity and the book-to-market ratio used in constructing $HML$ is affected by stock liquidity (Fang et al., 2009). Panel B2 includes all four FFC factors.

In Panels B1 and B2 of Table 1, $alpha$ is positive and highly significant for both periods, indicating a positive illiquidity premium adjusted for risk. In Panel B2,
alpha is 0.372% with $t = 3.80$ for Period I and 0.403% with $t = 3.12$ for Period II.\footnote{When using Fama and French’s (2015) factors $SMB$, $HML$, $RMW$ and $CMA$ and the momentum factor $UMD$, I find that for Period I, alpha is 0.424 with $t = 3.86$ and for Period II alpha is 0.265 with $t = 2.02$. As to the significance of the illiquidity premium in the recent period, see supporting evidence in Huh (2014) for the period 1983–2009 which extends 12 years beyond Amihud’s (2002) test period.} The positive and highly significant coefficient of $SMB$ reflects the well-known positive relation between liquidity and size. The positive slope coefficient of $HML$ indicates the greater illiquidity of the stocks with high book-to-market

<table>
<thead>
<tr>
<th>Periods</th>
<th>I: 1964 to 1997</th>
<th>II: 1998 to 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Statistics of IML</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.635 (4.47)</td>
<td>0.430 (2.14)</td>
</tr>
<tr>
<td>Median</td>
<td>0.641</td>
<td>0.218</td>
</tr>
<tr>
<td>% positive</td>
<td>60.5% (4.24)</td>
<td>56.7% (2.08)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Panel B: Regression of IML on Risk Factors</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel B1: Regressions of IML on the Market Excess Return</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha</td>
<td>0.714 (5.18)</td>
<td>0.552 (2.86)</td>
</tr>
<tr>
<td>$RM_f$</td>
<td>$0.154$ ($-2.85$)</td>
<td>$-0.221$ ($-4.70$)</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.053</td>
<td>0.097</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Panel B2: Regressions of IML on Fama–French–Carhart factors</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>0.372 (3.80)</td>
<td>0.403 (3.12)</td>
</tr>
<tr>
<td>$RM_f$</td>
<td>$-0.264$ ($-8.92$)</td>
<td>$-0.317$ ($-9.14$)</td>
</tr>
<tr>
<td>SMB</td>
<td>0.687 (11.89)</td>
<td>0.676 (9.58)</td>
</tr>
<tr>
<td>HML</td>
<td>0.304 (7.44)</td>
<td>0.339 (6.62)</td>
</tr>
<tr>
<td>$MOM$</td>
<td>0.029 (1.01)</td>
<td>0.015 (0.51)</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.621</td>
<td>0.613</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Panel C: Model (1) with $JAN_t$</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>0.398 (3.82)</td>
<td>0.513 (3.70)</td>
</tr>
<tr>
<td>Four factors included</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$JAN$</td>
<td>0.442 (1.04)</td>
<td>$-1.183$ ($-3.25$)</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.624</td>
<td>0.616</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Panel D: Statistics for Rolling 1-Month-Ahead alpha$_t$</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.523 (5.73)</td>
<td>0.333 (2.58)</td>
</tr>
<tr>
<td>Median</td>
<td>0.480</td>
<td>0.363</td>
</tr>
<tr>
<td>% positive</td>
<td>62.7% (5.05)</td>
<td>57.9% (2.45)</td>
</tr>
<tr>
<td>$N$</td>
<td>408</td>
<td>240</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Panel E: Regression of 1-Month-Ahead alpha$_t$ on a constant and $JAN_t$</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.489 (5.13)</td>
<td>0.439 (3.30)</td>
</tr>
<tr>
<td>$JAN$</td>
<td>0.412 (1.25)</td>
<td>$-1.274$ ($-2.76$)</td>
</tr>
</tbody>
</table>

Table 1: Estimates of Risk-Adjusted Returns on an Illiquid-Minus-Liquid (IML) Portfolio.
Table 1: Continued.

Description: The variable IMLt is the return on an illiquid-minus-liquid portfolio for month t, the differential return between the highest and lowest quintile portfolios of stocks sorted on their illiquidity, measured by the average daily value of \( \text{ILLIQ} = \frac{|\text{return}|}{\text{dollar volume}} \). In November of each year, stocks are sorted into three portfolios by SD, the standard deviation of their daily returns and, within each tercile portfolio, stocks are sorted into five portfolios by their \( \text{ILLIQ} \) value, producing 15 (3 \( \times \) 5) portfolios. The variables \( \text{ILLIQ} \) and SD are calculated over 12 months. For each portfolio, the value-weighted average return is calculated for each month \( t \) from January to December of the following year, using \( \text{ILLIQ} \) and SD for the month of November of the previous year. The variable IMLt is the average return on the three highest \( \text{ILLIQ} \) quintile portfolios (across volatility portfolios) minus the average return on the three lowest \( \text{ILLIQ} \) quintile portfolios. We use NYSE/AMEX stocks and apply some filters. The returns are in monthly percentage points. Estimations are carried out for the period 1964 to 1997, as in Amihud (2002), and for the years that follow, 1998 to 2017. The \( t \)-statistics of the estimated coefficients employ robust standard errors (White, 1980). In parentheses next to “\% positive” are values from a \( z \)-test approximation of the binomial test of the proportion against the null of 50%, the chance result.

(Panel A) shows the statistics for IMLt. (Panel B) presents \( \alpha \) and the \( \beta \) coefficients of the FFC factors obtained from the regression model

\[
\text{IML}_t = \alpha + \beta_{\text{RMrf}} \cdot \text{RMrf}_t + \beta_{\text{SMB}} \cdot \text{SMB}_t + \beta_{\text{HML}} \cdot \text{HML}_t + \beta_{\text{UMD}} \cdot \text{UMD}_t + \epsilon_t
\]  

(1)

where \( \text{RMrf} \) is the market excess return over the risk-free rate, \( \text{SMB} \) and \( \text{HML} \) are the Fama and French (1993) factors of size and the book-to-market (BE/ME) ratio, and \( \text{UMD} \) is the Carhart (1997) momentum factor. (Panel C) shows the estimations of Model (1) with an added variable, JANt, which equals one in the month of January and zero otherwise. (Panel D) presents the out-of-sample, one-month-ahead rolling \( \alpha \) values. Model (1) is estimated over a rolling window of 60 months beginning in January 1950. For month 61,

\[
\alpha_t = \text{IML}_t - \left[ \beta_{\text{RMrf},t-1} \cdot \text{RMrf}_t + \beta_{\text{SMB},t-1} \cdot \text{SMB}_t + \beta_{\text{HML},t-1} \cdot \text{HML}_t + \beta_{\text{UMD},t-1} \cdot \text{UMD}_t \right]
\]

using the \( \beta \) values estimated from the preceding 60-month window. (Panel E) shows the results from a regression of \( \alpha_t \) on a constant and JANt.

Interpretation: The results in Panel A show that the average illiquidity premium is positive and significant.

The risk-adjusted illiquidity premium, estimated by \( \alpha \), is positive and significant. Panel B1 controls for market risk and Panel B2 controls for the four Fama–French–Carhart risk factors. Panel D shows that the one-step-ahead risk-adjusted illiquidity premium is positive and significant, being approximately 4% per annum in period II.

Panels C and E control for the January effects. The risk-adjusted illiquidity premium is positive and significant for the entire year in Period I and for February to December in Period II.

ratio which is consistent with the evidence of Fang et al. (2009) on the negative effect of illiquidity on the market-to-book ratio. The momentum factor is insignificant.

Panel C of Table 1 presents the estimation results of Model (1) with an added dummy variable, JANt, which equals one in the month of January and zero otherwise. This model tests if the illiquidity effect is confined to the month of January, as is the case with the small firm effect. The results show that, in Period I, the January effect is positive and insignificant but, in Period II, it becomes negative.
and significant. Thus, in the recent 20-year period, the average illiquidity premium is positive only for the 11 months from February to December.

Panel D of the table presents estimates of 1-month-ahead out-of-sample estimates of \( \alpha_t \) of \( IML_t \). I first estimate the coefficients of Model (1) over a 60-month rolling window that ends in month \( t - 1 \). Then, the estimated coefficients \( \beta_{K,t-1} \) for \( K = RMrf, SMB, HML, \) and \( UMD \) are used to calculate \( \alpha_t \), conditional on the realized factor returns in month \( t \):

\[
\alpha_t = IML_t - [\beta_{RMrf,t-1} \times RMrf_t + \beta_{SMB,t-1} \times SMB_t + \beta_{HML,t-1} \times HML_t + \beta_{UMD,t-1} \times UMD_t].
\]

This procedure is repeated by rolling forward the 60-month estimation window 1 month at a time.

Panel D presents statistics of the series \( \alpha_t \) for the period January 1964 to December 2017. The mean of out-of-sample \( \alpha_t \) for Period I is 0.523%, with \( t = 5.73 \), the median is 0.480%, which is close to the mean, and the fraction of \( \alpha_t > 0 \) is 0.627, significantly greater than 0.50 which is the chance result. For Period II—after Amihud's (2002) analysis period—the mean \( \alpha_t \) is 0.333%, with \( t = 2.58 \), the median is 0.363%, again close to the mean, and the fraction of \( \alpha_t > 0 \) is 0.579, which is significantly greater than 0.50.

Panel E of Table 1 presents the test results of the January effect on \( \alpha_t \). In Period I, the January effect is positive and in Period II it is negative. For the month of January alone, the means \( \alpha_t \) in Period I \( (n = 34) \) and Period II \( (n = 20) \) are, respectively, 0.901% \( (t = 1.76) \) and −0.834% \( (t = −2.25) \). The flip in the sign of the mean \( \alpha_t \) in January is puzzling. For the 11 months from February to December, the means of \( \alpha_t \) are positive and significant in both periods and of similar magnitude.

Figure 1 plots the 12-month moving average of \( \alpha_t \) for a longer sample period, 1955 to 2017. The series is mostly in positive territory, including in recent years. Its most negative value is in 2000, the year when the dot-com bubble burst. Over the entire period the mean \( \alpha_t \) is 0.475% with \( t = 7.01 \), the median is 0.480% which is close to the mean and the proportion of \( \alpha_t > 0 \) is 0.627 which is significantly higher than 0.50, the chance result.

4 The Effects of the Time-Series of Market Illiquidity Shocks on \( IML \)

Amihud (2002) finds a negative and significant relation between market illiquidity shocks and realized stock returns. A positive shock to illiquidity, which is highly persistent, raises the expected level of illiquidity and makes investors demand higher expected returns on stocks. Consequently, stock prices fall to raise expected returns, assuming that cash flows are unaffected by market illiquidity shocks.\(^{13}\)

\(^{13}\)See a formal derivation in Acharya and Pedersen (2005).
Figure 1: 12-Month Moving Average of 1-Month-Ahead Rolling $\alpha_t$.

**Description:** This figure plots a 12-month moving average of the monthly 1-month-ahead $\alpha_t$, calculated as

$$\alpha_t = \text{IML}_t - [\beta_{RMf,t-1} \cdot \text{RMf}_t + \beta_{SMB,t-1} \cdot \text{SMB}_t + \beta_{HML,t-1} \cdot \text{HML}_t + \beta_{UMD,t-1} \cdot \text{UMD}_t].$$

where $\text{IML}_t$ is the monthly return on an illiquid-minus-liquid portfolio (see Table 1 for details) and the $\beta$ values are estimated over 60 months preceding month $t$ from the regression Model (1). The sample period is 1964 to 2017. The numbers on the $y$-axis are monthly returns as a percentage.

**Interpretation:** The monthly risk-adjusted illiquidity premium is generally positive. It is low in the 1990s and deeply negative in the late 1990s, and it recovers to become mainly positive thereafter.

It follows that the market return, which is the sum of expected and unexpected returns, is negatively affected by contemporaneous illiquidity shocks and positively affected by lagged illiquidity, which is a proxy for expected illiquidity. The analysis follows Merton’s (1980) analysis of expected market returns as an increasing function of market volatility, causing expected returns to change through time and the analysis of French *et al.* (1987), who find that the market return is a negative function of unexpected market volatility and a positive function of expected volatility.

Earlier studies show that stock prices are negatively impacted by changes in illiquidity. Amihud *et al.* (1990) study the stock market crash of October 19, 1997, when the S&P 500 share index fell by more than 20% and illiquidity sharply increased. The average quoted bid–ask spread (in dollars) increased by 63% relative to its average level in the first week of October and there was a sharp decline in market depth, the size of orders that can be exercised at the quoted bid and ask prices. This study proposes that the price decline occurred partly because of investors’ recognition that illiquidity is hurt by program trading, which was prevalent at the time and not as high as previously thought. The finding is
that, across firms on the day of the crash, stock returns were negatively related to changes in illiquidity. Amihud et al. (1997) resolve the issue of causality by presenting evidence that prices rise on stocks that undergo an exogenous increase in liquidity. The Tel Aviv Stock Exchange gradually transferred stocks from trading in a once-a-day call auction session to more continuous trading sessions, which improved liquidity. Measures of stock liquidity—Amivest and trading volume—rose. This led to a sharp rise in the price of the transferred stocks. Muscarella and Piwowar (2001) find similar results for the Paris Bourse where prices increased for stocks whose liquidity improved when they were transferred from call trading to continuous trading. In their study of the Japanese market, Amihud et al. (1999) find that stock prices increased when companies reduced the minimum order size in their stocks which facilitated trading by small retail investors who are viewed as uninformed liquidity traders and consequently improved stock liquidity. Across stocks, price appreciation was an increasing function of the liquidity improvement that followed from this change.14

In addition to the price decline because of the rise in expected returns when expected illiquidity rises, I suggested that illiquid stocks suffer further price declines because of the flight to liquidity, where investors substitute away from illiquid into liquid assets when expected illiquidity rises.15 For liquid stocks that become more attractive as illiquidity rises, the two effects work in opposite directions and, therefore, the negative effect of illiquidity shocks on prices of liquid stocks is weaker. The differential negative impact of illiquidity shocks on illiquid stocks, defined as illiquidity systematic risk (or illiquidity beta), is used in the pricing of stocks and bonds by Pastor and Stambaugh (2003), Acharya and Pedersen (2005), Lin et al. (2011), and Bongaerts et al. (2018), among others.

In my 2002 paper, I studied the differential effect of illiquidity shocks on stock returns across five size-based stock portfolios. I now estimate the differential effect of illiquidity shocks on IML, the high-minus-low illiquidity quintile portfolio, with the following model:

\[
IML_t = a_0 + b_1 \times uMILLIQ_t + b_2 \times RMrf_t + b_3 \times JAN_t + u_t,
\]

where \(MILLIQ_t\) is the market illiquidity, calculated as the value-weighted monthly average of stock illiquidity, \(uMILLIQ_t\) is the 1-month-ahead unexpected illiquidity using an AR(2) model of \(\log(MILLIQ_t)\) with a trend variable that is estimated dynamically over a lagged rolling window of 60 months. \(RMrf_t\) is added as a control variable given the contemporaneous negative relation between illiquidity shocks and the market excess return.

14 In a study of the Tel Aviv Stock Exchange, Amihud et al. (2003) find a significant increase in stock prices subsequent to the exercise of warrants that significantly improved stock liquidity by consolidating trading of two similar securities into one.

15 Supporting evidence for such a pattern is presented by Huang (2010) who finds that mutual funds switch from illiquid to liquid holdings when they anticipate adverse market conditions. Acharya et al. (2013) find evidence of a flight to liquidity in corporate bonds that is distinct from flight to safety.
Table 2: Effect of Illiquidity Shocks on IML.

Description: This table presents estimation results of the model

\[ IML_t = a_0 + b_1 \cdot uMILLIQ_t + b_2 \cdot RMrf_t + b_3 \cdot JAN_t + \epsilon_t, \]  

where \( uMILLIQ_t \) is the 1-month-ahead unexpected market illiquidity where \( MILLIQ_t \) is market illiquidity, the value-weighted monthly average of stock illiquidity. Unexpected illiquidity is the difference between \( \log MILLIQ_t \) and its predicted value using an AR(2) model of \( \log MILLIQ_t \) with a trend variable that is estimated over a rolling window of 60 months up to month \( t - 1 \), and whose coefficients are used to obtain a predicted value of \( \log MILLIQ_t \). The \( t \)-statistics in parentheses employ robust standard errors.

Interpretation: The table shows that illiquidity shocks, measured by \( uMILLIQ_t \), have a more negative effect on illiquid stocks than on liquid stocks.

The monthly market illiquidity series \( MILLIQ_t \) is constructed by averaging the market’s daily cross-stock average illiquidity over the days of month \( t \). The daily average illiquidity for each day \( d \) of month \( t \) is a weighted average of the daily values of \( ILLIQ_{j,d,t} \) using all stocks \( j \) that are employed in the construction of \( IML_t \). Then, I calculate the illiquidity shocks as follows. Over a window of 60 months, I conduct a regression of the logarithm of \( MILLIQ_t \) on its two lags, a trend line and a constant.\(^{16}\) Finally, I calculate \( uMILLIQ_t \) for month 61 as the difference between \( \log MILLIQ_t \) and its predicted value, using the estimated coefficients from the preceding 60 months. I then roll the window ahead by one month and repeat the procedure. Thus, there is no hindsight in the generation of \( uMILLIQ_t \).

Table 2 presents the estimation results of Model (2) for both sample periods. The research question is whether the effect of illiquidity shocks is greater for less liquid stocks, as proposed in Amihud (2002). The estimations of Model (2), presented in Table 2 show that market illiquidity shocks negatively affect \( IML_t \) in both periods. The negative effect of \( uMILLIQ_t \) is statistically significant in both periods when \( RMrf_t \) is included. Notably, when \( RMrf_t \) itself is regressed on \( uMILLIQ_t \), the slope coefficient is negative and highly significant in both periods. Similarly, when returns on the highest and lowest illiquidity quintile portfolios which are used to construct \( IML \) are separately regressed on \( uMILLIQ_t \), the coefficients are negative.

\(^{16}\) The results through this section are qualitatively similar when the autoregressive coefficients are adjusted by the bias correction method of Shaman and Stine (1988, 1989) that accounts for finite-sample bias.
and significant in both periods regardless of whether $RMrf_t$ is included in the model. The January effect is positive in Period I and negative in Period II as it is in the cross-section analysis.

I add $\log\text{MILLIQ}_{t-2}$, an estimator of expected illiquidity, to the models in columns (2) and (4), to test Amihud’s (2002) proposition that the effect of expected market illiquidity is greater for more illiquid stocks. I find that its coefficient of $\log\text{MILLIQ}_{t-2}$ is insignificantly different from zero. In Period I it is 0.177 with $t = 1.11$ and in Period II it is $-0.089$ with $t = -0.43$.

In summary, there is evidence on a significant and more negative effect of illiquidity shocks on the returns of less liquid stocks but there is no support for the hypothesis on the differential effect of expected illiquidity on the expected returns between illiquid and liquid stocks. The difficulty of finding a significant positive effect of expected illiquidity on expected returns in time series is similar to the difficulty of finding a significant positive effect of expected risk on expected returns, noted in French et al. (1987). Guo and Whitelaw (2006) discuss the econometric problems in estimating the effect of expected volatility on ex ante return. Some of these problems apply for the estimation of the effect of expected illiquidity on expected return. Further study may be needed along the line of Guo and Whitelaw's (2006) study of Merton's (1980) prediction of the varying effect of expected volatility on expected returns, in which they estimate a conditional expectation model based on state variables. The analysis needs to account not only for the effect of macroeconomic state variables on expected illiquidity but also for their effect on the price or premium of illiquidity. Brunnermeier and Pedersen (2009) show theoretically that both the level and the price of market liquidity vary over time as a function of funding liquidity and Acharya et al. (2013) find that illiquidity is priced in the time-series of corporate bond returns mainly in periods of financial distress.

5 Concluding Remarks

Amihud and Mendelson (1986) proposed that investors require a return premium as compensation for illiquidity costs. They predicted that across stocks, illiquidity affects expected return, and they presented evidence that supports this prediction. Amihud (2002) continued this line of research by showing the time-series effect of illiquidity on stock returns.

Amihud and Mendelson’s (1986) theory proposes that the illiquidity premium differs by investors’ holding-period clientele and exceeds expected illiquidity costs because of investors’ funding constraints. The theory predicts a positive relation between expected returns and illiquidity costs. This is the important takeaway; the specific measure of illiquidity that is used in empirical tests is secondary.

Illiquidity has a number of dimensions that are hard to capture in a single measure, including fixed costs, variable costs—price impact costs that increase in
the traded quantity—and opportunity costs. The variable ILLIQ is one proxy, just as the bid–ask spread was a proxy in Amihud and Mendelson (1986).\textsuperscript{17} ILLIQ has been shown to have a high positive correlation with high-frequency measures of illiquidity (estimated from intraday data) but it does not capture everything about illiquidity. It is possible to combine ILLIQ and other low-frequency measures of illiquidity into one measure using principal component analysis as do Korajczyk and Sadka (2008) for high-frequency measures of liquidity. Among low-frequency measures of illiquidity are those of Roll (1984), Lesmond \textit{et al.} (1999), Brennan \textit{et al.} (1998), Chordia \textit{et al.} (2001), Pastor and Stambaugh (2003), Liu (2006), Hasbrouck (2009), Goyenko \textit{et al.} (2009), Holden (2009), Das and Hanouna (2010), Corwin and Schultz (2012), Fong \textit{et al.} (2017), and Abdi and Ronaldo (2017). Some illiquidity measures employ volume and volatility. Amihud (2002, p. 34) points out that ILLIQ is “strongly related to the liquidity ratio known as the Amivest measure, the ratio of the sum of the daily volume to the sum of the absolute return.” Amivest, used in Amihud \textit{et al.} (1997, 1999), is defined as “a liquidity measure that calculates the dollar value of trading that would occur if prices changed 1%.”\textsuperscript{18} Harris and Amato (2018) test the pricing of a set of illiquidity measures that are based on daily return and volume data. They use Amivest\textsuperscript{-1} and the invariance illiquidity measure of Kyle and Obizhaeva (2016), the third root of the variance of returns to the average dollar volume. This measure is approximately the product of Amivest\textsuperscript{-1/3} and the third root of the return standard deviation which is closely related to average absolute return. Harris and Amato (2018) find that these alternative liquidity measures significantly predict expected returns.

The magnitude of the effect of illiquidity on expected return varies over time. In Amihud (2002) the effect of ILLIQ on expected return is lower in the second half of the sample period than in the first half. Similar evidence is presented in Ben-Rephael \textit{et al.} (2015), Drienko \textit{et al.} (2018), and Harris and Amato (2018). The analysis in Table 1, which employs a different methodology than that used in the aforementioned studies, shows lower risk-adjusted illiquidity premium in the recent period that is still positive and significant. Figure 1 shows considerable variations over time in the average illiquidity premium. In general, long-term trends in the effect of illiquidity on expected return are partly affected by institutional changes in the securities markets which affect liquidity and trading and by the means developed to circumvent the costs of illiquidity. Since the 1980s we observe a strong downward trend in market ILLIQ. There may be a number of reasons for that. Since the 1980s there has been a strong entry of discount brokers which increased competition in the brokerage industry, facilitated trading, and lowered the cost of trading especially for small investors. Amihud \textit{et al.} (1999) show that

\textsuperscript{17}Trading volume data were not publicly available when Amihud and Mendelson’s (1986) study was written.

\textsuperscript{18}See http://www.yourdictionary.com/amivest-liquidity-ratio, quoted from Webster New World Finance and Investment Dictionary.
facilitating trading increases the presence of small investors in the market and improves stock liquidity. The reduction in the minimum quoted tick size from $1/8 to $1/16 in 1997 and the further reduction to $0.01 in 2001 helped increase market liquidity\textsuperscript{19} and reduce the illiquidity premium. Recent developments in automated trading and trading schemes to reduce trading costs further improve market liquidity and reduce the illiquidity premium.

Mutual funds and index funds enabled effective trading in securities—even less liquid ones—while saving in trading costs. At the end of a trading day a typical fund offsets buy and sell orders for its units thus enabling its investors to transfer ownership of the underlying securities without having to trade these securities. In this way, the fund saves the cost that investors would have incurred if they had directly bought and sold the securities that constitute the fund portfolio. The fund needs to trade only to the extent required by the residual unmatched demand for its units.

Another way to save on trading cost is by enabling trading in a liquid security that represents a claim on a portfolio of illiquid assets such as securitized loans, mortgage backed securities, and exchange traded funds. Unlike risk which can be reduced through portfolio diversification illiquidity is not reduced when holding a portfolio of illiquid stocks. Illiquidity is additive: buying and selling a portfolio of illiquid assets entails bearing the sum of the illiquidity costs of its components.\textsuperscript{20} However, if a security that represents a claim on a portfolio of illiquid assets is liquid, its price will reflect a lower illiquidity premium than that of the underlying assets. This will permeate to the underlying securities, raise their price and reduce their illiquidity premium. This idea has been presented by Amihud and Mendelson (1988) in the context of securitization of loans. Banks and financial firms pool and package individual loans, which are highly illiquid, into standard debt securities that are liquid traded assets. Competition between financial intermediaries passes the benefits of increased liquidity to the borrowers in the form of lower interest rates or a lower illiquidity premium.

Similarly, the market has developed exchange-traded funds (ETFs) that are often more liquid than their constituent securities that include illiquid stocks and bonds. As in Amihud and Mendelson’s (1988) analysis of securitization, ETFs transfer the benefit of their higher liquidity to their constituent securities in the form of a lower illiquidity premium. The existence of liquid ETFs partly underscores the importance of illiquidity costs. Ben-Rephael et al. (2015) propose that the expansion of investment through ETFs explains the decline in the illiquidity premium that they document.

However, there is no free lunch. A fund enables saving in trading costs while making its investors face an opportunity cost since they are restricted to a given portfolio which could deviate from what they consider optimal. Investors face a

\textsuperscript{19}See Bessembinder (2003) on the improvement in liquidity following the 2001 decimalization.
\textsuperscript{20}Amihud and Mendelson (2010).
tradeoff: on the one hand, they incur higher trading costs when holding directly their optimal portfolio of securities; on the other hand, they bear the cost of deviating from optimality and losing flexibility when holding a liquid fund whose return is imperfectly correlated with what they consider as their optimal portfolio. In addition, a fund charges a continuous management fee, which effectively shifts part of the cost from trading costs to these fees. Thus, while mutual funds, ETFs, and similar investment vehicles reduce the cost of illiquidity, which is what they are designed to do, they do not eliminate them.\textsuperscript{21}

Recent developments in capital markets raise the value of liquidity for some investors. Liquid securities are in greater demand by high-frequency traders, consistent with Amihud and Mendelson’s (1986) prediction that the demand for assets with different illiquidity costs is affected by investor clienteles that differ in their expected holding periods. Increased trading frequency makes liquidity more valuable. Another development is the rise of activist investors, who favor investing in firms with liquid stock (Fos, 2017).

Ultimately, the evidence supports the proposition that illiquidity is priced, using a variety of proxies, and that there is a positive illiquidity premium which varies over time. Amihud et al. (2005, 2013) review evidence on the pricing of illiquidity and illiquidity risk. Drienko et al. (2018) and Harris and Amato (2018) find that the cross-sectional effect of illiquidity is positive and significant for their entire estimation period although it has diminished recently. Worldwide, Amihud et al. (2015) find that stock illiquidity has a positive effect on expected returns.

In conclusion, there are now two major theories of asset pricing. One is based on asset risk and the other on asset illiquidity. Both theories propose that these asset characteristics, which are undesirable by investors, entail an expected return premium as compensation. Empirical evidence on the cross-section determinants of expected return shows that the pricing of illiquidity is no less—and perhaps more—robust than the pricing of risk.\textsuperscript{22} The existence of a positive risk premium is supported by evidence on a positive and significant mean of $R_{Mrf}$, the excess return of risky stocks over the risk-free rate. Similarly, there is a positive (risk-adjusted) mean of $IML$, the excess return of illiquid over liquid stocks. Although in some periods we observe that the average realized market risk premium $R_{Mrf}$,

\textsuperscript{21}Petajisto (2016) finds that deviations of ETF prices from their net asset value reflect the cost of redeeming and creating units of ETFs. These deviations are greater in funds of illiquid securities.

\textsuperscript{22}There is meager evidence on the significant pricing of stock systematic ($\beta$) risk. However Bali and Engle (2010) find that the dynamically conditional covariance risk is positively priced. Evidence on a negative effect of risk, measured by return standard deviation, on the cross section of expected stock returns is presented in Amihud (2002), Drienko et al. (2018), and Harris and Amato (2018). Ang et al. (2006, 2009) find idiosyncratic risk to have a negative and significant effect on the cross section of expected returns. However, Ghysels et al. (2005) find a significant positive relation between risk and expected return and Han and Lesmond (2011) suggest that the negative relation between idiosyncratic risk and expected return is due to idiosyncratic biases in estimated returns. Empirical evidence on the determinants of security returns is reviewed in Bodie et al. (2018, Ch. 13).
is insignificantly different from zero,\(^\text{23}\) we do not conclude from this evidence that investors are not averse to risk nor do we dismiss the existence of a risk premium. Similarly, evidence that in some subperiods there is insignificant pricing of illiquidity does not imply that investors are not averse to illiquidity cost nor should it dismiss the existence of an illiquidity premium.

References


\(^{23}\)With hindsight, we observe a number of periods of 10 years with a negative average monthly RMrf and there are a number of periods of 20 years with positive but insignificant average monthly RMrf. There is a period of 40 years, 1969 to 2008, for which the average RMrf is 0.338% with \(t = 1.61\), insignificant.


