Journal of Financial Economics 130 (2018) 48–73

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Micro(structure) before macro? The predictive power of aggregate illiquidity for stock returns and economic activity[☆]

Yong Chen[a,*], Gregory W. Eaton[b], Bradley S. Paye[c]

[a] Mays Business School, Texas A&M University, 360J Wehner Building, College Station, TX 77843, USA
[b] Spears School of Business, Oklahoma State University, 333 Business Building, Stillwater, OK 74078, USA
[c] Pamplin College of Business, Virginia Tech, 1016 Pamplin Hall, Blacksburg, VA 24061, USA

ARTICLE INFO

Article history:
Received 9 February 2016
Revised 1 March 2017
Accepted 31 March 2017
Available online 30 May 2018

JEL classification:
G1
C13

Keywords:
Stock market liquidity
Stock return predictability
Macroeconomic forecasts
Transaction costs
Equity premium

ABSTRACT

This paper constructs and analyzes various measures of trading costs in US equity markets covering the period 1926–2015. These measures contain statistically and economically significant predictive signals for stock market returns and real economic activity. We decompose illiquidity proxies into a component capturing aggregate volatility and a residual. The predictive content of these components differs in important ways. Specifically, we find strong evidence that the component of illiquidity uncorrelated with volatility forecasts stock market returns. Both the volatility and residual components of illiquidity contain information regarding future economic activity.

1. Introduction

Liquidity conditions fluctuate over time in securities markets. In addition, Chordia et al. (2000), Hasbrouck and Seppi (2001), and Huberman and Halka (2001) document a significant common component in the time-series variation of liquidity.[1] Fluctuations in aggregate liquidity potentially

☆ We are extremely grateful to a number of researchers who provided data and code used in this paper. Specifically, Ekkehart Boehmer and Julie Wu furnished high-frequency effective spread measures for NYSE firms and associated code. Charles Trzcinka, Shane Corwin, and Paul Schultz provided code used to compute several liquidity proxies. Bernt Arne Ødegaard furnished code and details regarding liquidity proxies used in Næs et al. (2011). We have benefited from helpful comments of Yakov Amihud, Wayne Ferson, Charles Jones, Bob Korajczyk, Bernt Arne Ødegaard (EFA discussant), Matthew Pritsker (MFA discussant), Pedro Saffi (FIRS discussant), Avanidhar Subrahmanyam, Guofu Zhou, and, in particular, the anonymous referee, as well as seminar and conference participants at Texas A&M University, the University of Georgia, the June 2014 CREATES Econometrics Workshop, the 2015 European Finance Association (EFA) Meetings, the 2015 Financial Management Association (FMA) Meetings, the 2015 Financial Econometrics: Challenges and Directions for Future Research Conference, the 2016 Financial Intermediation Research Society (FIRS) Meeting, and the 2016 Midwest Financial Association (MFA) Meetings. Chen acknowledges research support from the Mays Business School at Texas A&M University. Paye acknowledges support from CREATES-Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

* Corresponding author.

E-mail addresses: ychen@mays.tamu.edu (Y. Chen), gregory.eaton@okstate.edu (G.W. Eaton), bpaye@vt.edu (B.S. Paye).

[1] Various studies extend evidence of commonality in liquidity beyond US equity markets. Chordia et al. (2005) find evidence of commonality in aggregate spreads and depth across US equity and Treasury bond markets. Brockman et al. (2009) and Karolyi et al. (2012) document commonality in liquidity across countries and exchanges. Subrahmanyam (2007) finds commonality between liquidity in equity markets and liquidity in real estate investment trusts (REITs) markets. Mancini et al. (2013) identify common variation in liquidity across a number of asset classes and markets, including major foreign exchange currency pairs, US equities, and US government and corporate bond markets.

https://doi.org/10.1016/j.jfineco.2018.05.011
0304-405X/

constitute a priced risk factor for stocks and other assets (Acharya and Pedersen, 2005), and numerous papers provide empirical evidence concerning the pricing of liquidity risk.

The implications of time-varying aggregate liquidity range beyond the cross-sectional pattern of security returns. Theoretical models link aggregate liquidity with time-variation in the equity premium. In the model of Acharya and Pedersen (2005), for example, persistent variation in liquidity implies that liquidity forecasts future stock returns. A related theoretical literature emphasizes connections between financial frictions and macroeconomic activity. Illiquidity, coupled with other financial frictions, can generate nonlinear amplification effects and exacerbate economic downturns (Brunnermeier and Pedersen, 2009). This theoretical literature suggests that empirical measures of aggregate illiquidity should contain predictive signals for future stock returns and economic activity.

Empirically, Amihud (2002) and Jones (2002) find that certain aggregate liquidity measures forecast US stock market returns. However, relatively few subsequent papers emphasize such measures as return forecasting variables. For example, Goyal and Welch (2008) document the in- and out-of-sample forecasting power associated with a large set of return forecasting variables, but none of their variables explicitly measure aggregate liquidity. Bekaert et al. (2007) find that a liquidity proxy based on the proportion of zero return days forecasts stock returns in emerging markets, but it is unclear whether similar results obtain for developed markets. Næs, Skjeltorp and Ødegaard (2011) find that variation in aggregate liquidity predicts macroeconomic outcomes. Chen et al. (2016) report that certain illiquidity measures predict the onset of recessions in US data.[2]

This paper provides new, relatively comprehensive evidence concerning the predictive content of aggregate illiquidity for stock returns and real economic activity. We construct a variety of measures of market frictions at monthly and quarterly frequencies using Center for Research in Security Prices (CRSP) daily data for stocks listed on the New York Stock Exchange (NYSE).[3] These time series cover a 90-year period from 1926–2015 and include multiple proxies for the effective spread, several alternative measures of price impact, and a measure of aggregate price delay. Since our measures relate positively to transaction costs, we refer to them as "illiquidity" measures.

Consistent with earlier evidence in Jones (2002) for bid-ask spreads on Dow Jones stocks, aggregate illiquidity contains a cyclical component. Liquidity tends to erode during economic recessions, and markets are considerably more illiquid during both the Great Depression and the recent financial crisis. We propose two key adjustments to conventional aggregate illiquidity measures. The first adjusts for an embedded market volatility component. The second adjusts for structural breaks in illiquidity proxies, particularly those connected with tick-size reductions occurring in 1997 and 2001. We illustrate the empirical importance of these adjustments by examining the predictive content of adjusted versus unadjusted illiquidity measures for excess stock market returns and various measures of economic activity.

Most aggregate illiquidity measures are positively correlated with stock market volatility. The connection between measures of market frictions and measures of volatility runs deep: we show that several common proxies for the effective spread are mechanically log-linear in the standard deviation of firm returns for the corresponding period. In a similar fashion, price impact measures of the form proposed by Amihud (2002) are approximately log-linear in return volatility. Although firm-level return volatility incorporates both idiosyncratic and systematic risk, cross-sectional averages of illiquidity proxies embed a component that is highly correlated with standard measures of market volatility. The embedded volatility component is concerning, because effects actually driven by market volatility may be falsely attributed to market illiquidity. For example, the intertemporal capital asset pricing model (ICAPM) of Merton (1973) implies a positive relationship between the conditional market risk premium and conditional volatility *even in the absence of market frictions*. Similarly, market volatility may relate to future economic activity and cross-sectional asset prices for reasons unrelated to liquidity conditions (see, e.g., Fornari and Mele, 2013; and Bansal et al., 2014, respectively).

Motivated by these observations, we decompose illiquidity measures into a component reflecting aggregate volatility and a residual. Importantly, we show that a simple measure of illiquidity based on the frequency of zero return days (Lesmond et al., 1999) can be interpreted as a spread measure that is already volatility-adjusted. After extracting a volatility component, other common spread measures are highly correlated with the zeros measure in the aggregate.

The second adjustment we make corrects illiquidity measures for structural shifts related to NYSE minimum tick-size reductions in 1997 and 2001. A number of aggregate illiquidity measures shift downward around these reductions. Effects are particularly sharp for the zeros liquidity proxy and several spread measures. Existing evidence regarding the impact of tick-size reductions on market liquidity is mixed. This evidence indicates that tick-size reductions lead to decreases in quoted and effective spreads, but also to associated reductions in quoted depth. Consequently, the net effect on transaction costs is unclear and potentially varies across alternative classes of investors.[4]

[2] Other studies linking illiquidity with macroeconomic activity focus on international settings (see, e.g., Smimou, 2014; Rai, 2015; and Smimou and Khallouli, 2015) and extensions to nonlinear specifications (see, e.g., Florackis et al., 2014). Beber et al. (2011) and Kayacetin and Kaul (2009) analyze the predictive content in aggregate equity order flow for the state of the economy. Cao et al. (2013) find that, in the context of hedge funds, liquidity timing is associated with better fund performance, suggesting positive investment value.

[3] Our convention of basing measures on NYSE stocks follows a number of papers in the literature, including Amihud (2002), Chordia et al. (2001), Goyenko et al. (2009), and Næs et al. (2011).

[4] Studies documenting lower effective spreads following tick-size reductions include Bacidore et al. (2003), Bessembinder (2003), Chakravarty et al. (2004), Goldstein and Kavajecz (2000), Jones and Lipson (2001), and Van Ness et al. (2000). The aggregate effective spread and several low-frequency proxies continue to fall for several years following decimaliza-

For example, Jones and Lipson (2001) find that total trading costs for institutional investors executing large trades increased following the 1997 reduction. Stark decreases in aggregate spread proxies around tick-size reductions likely overstate any true liquidity improvements. Consequently, we adjust illiquidity measures for breaks using several alternative approaches, including 'real time' approaches that mitigate concerns regarding spurious forecasting power attributable to look-ahead bias.

The remainder of the paper analyzes the predictive content associated with various aggregate illiquidity measures, with a focus on assessing the predictive power of the volatility component embedded in such measures relative to the residual component. We test whether illiquidity measures and their components forecast stock market returns and various measures of economic activity, including the growth rate in industrial production (or gross domestic product (GDP) growth and its components at the quarterly frequency) and the growth rate in unemployment.

We find little evidence that unadjusted illiquidity measures forecast stock returns. However, weak forecasting performance is potentially attributable to level shifts in illiquidity. Indeed, we obtain stronger evidence of stock return forecasting power for break-adjusted measures. This finding resembles results in Lettau and Van Nieuwerburgh (2008), who show that adjusting for level shifts in financial ratios produces stronger and more stable evidence of stock return predictability. The zeros measure of transaction costs proposed by Lesmond et al. (1999) significantly forecasts returns and performs best among break-adjusted measures. The zeros proxy is essentially uncorrelated with market volatility. This suggests that more consistent evidence of stock return forecasting power will obtain for volatility-adjusted illiquidity measures. Consistent with this intuition, we find that many break- and volatility-adjusted measures significantly forecast returns.

The intuition behind the importance of volatility-adjustment in stock return forecasting regressions is as follows. The return forecasting coefficient associated with unadjusted illiquidity measures is a weighted average of the forecasting coefficient for the volatility component and that for the volatility-adjusted (residual) component. In the data, the stock return forecasting coefficient associated with lagged aggregate volatility is negative.[5] As a result, estimates of stock return forecasting coefficients for raw illiquidity measures are often close to zero and insignificant, whereas slope estimates for volatility-adjusted measures are positive and significant. It is important to emphasize that our results clarify the predictive content in illiquidity measures, but do not necessarily imply a negative risk-return relation. For example, Ghysels et al. (2005) show that alternative volatility models that incorporate a longer window length in forecasting variance yield a significant positive relation.

The predictive power associated with several break- and volatility-adjusted measures is economically significant. One-month-ahead predictive regressions for excess returns based on these measures yield in-sample R^2-values of approximately 0.5–1.5%. This compares favorably to R^2-values for many traditional predictors over our sample period, including the short-term interest rate, term spread, default spread, and popular financial ratios such as the book-to-market ratio and earnings-to-price ratio. Break- and volatility-adjusted illiquidity measures remain statistically significant predictors of stock returns in multivariate predictive regressions that include an array of traditional return forecasting variables. Out-of-sample return forecasting results confirm in-sample evidence of predictive power, and the adjusted illiquidity measures generally outperform both the historical average and other common return forecasting variables in out-of-sample settings.

Because illiquidity measures exhibit long memory and illiquidity shocks are contemporaneously (negatively) correlated with return shocks, econometric concerns arise related to finite sample bias and related inference problems (Stambaugh, 1999). Key results regarding return predictability continue to obtain under alternative bootstrap inference methods that accommodate these features of the data. Slope coefficients remain relatively stable across subsamples and results are robust to alternative break-adjustment procedures, including real time approaches. Finally, we document that the predictive power of adjusted illiquidity measures concentrates in recessions and financial crisis periods, including the mid-1970s and around the recent financial crisis.

Turning to measures of economic activity, we find strong evidence that aggregate liquidity predicts economic activity, controlling for past activity, corroborating Næs et al. (2011). Consistent with intuition, more illiquid markets predict lower future output growth and higher future unemployment. Adjusting for structural shifts generally improves forecasting performance. Most series that exhibit a downward break following tick-size reductions only predict economic activity after break adjustment. Upon decomposing illiquidity measures into a volatility component and a residual, we find that a portion of illiquidity measures' predictive power for economic activity derives from the volatility component.[6] Although the residual component of illiquidity possesses weak predictive power in univariate models, conditioning on a wider information set produces stronger evidence of marginal predictive power for volatility-adjusted measures. Consequently, we conclude that both components of illiquidity are informative concerning future economic activity.

Quarterly data permit additional measures of economic activity and provide an additional robustness check for key stock return forecasting results. We obtain qualitatively

tion in 2001. A possible explanation for this pattern is the introduction of 'OpenBook' to the NYSE, which allowed subscribing market participants real time access to the limit order book. Boehmer et al. (2005) report evidence of reductions in the effective spread associated with this move toward ex ante limit order book transparency.

[5] This is consistent with French et al. (1987), who find an insignificant and sometimes negative risk-return relation using a volatility proxy based on the sum of squared daily return in the prior month (a measure highly correlated with our aggregate volatility measure).

[6] Our results concerning the predictive power of the volatility component of illiquidity measures are consistent with Campbell et al. (2001) and Fornari and Mele (2013), who find that volatility helps predict GDP growth. In related work, Giglio et al. (2016) find that systemic risk measures are able to predict macroeconomic downturns.

similar results concerning the forecasting power of adjusted illiquidity measures using quarterly data. In a final extension, we extract a principal component from break- and volatility-adjusted illiquidity measures that efficiently summarizes information across many illiquidity proxies. The corresponding principal component significantly forecasts excess stock market returns and most measures of economic activity.

Our paper contributes to a stream of literature addressing the time-series properties of aggregate liquidity. Chordia, Roll and Subrahmanyam (2001) document regularities in the time series of daily liquidity and trading activity and explore the causes of associated movements. Jones (2002), an important predecessor to our paper, documents the cyclical behavior of aggregate liquidity and provides early evidence regarding the predictive power of liquidity measures for US stock returns. We find that many illiquidity measures are subject to structural shifts, with important shifts occurring around tick-size reductions near the turn of the century, and demonstrate the importance of adjusting for shifts from a forecasting perspective. We also illustrate the importance of adjusting for volatility in uncovering the relation between illiquidity and the equity premium. In so doing, we extend evidence in Bekaert et al. (2007) to US data and develop an interpretation of the zeros proxy as a volatility-adjusted illiquidity measure. From a broader perspective, our results indicate that the measurement of aggregate illiquidity conditions is nontrivial due to structural instability and the potential for proxies to subtly embed a measure of stock market volatility. Our paper applies new, adjusted aggregate illiquidity measures in a forecasting context. A range of potential future applications is available, including, for example, application to the pricing of liquidity risk in the cross-section.

2. Data

This section describes our data. We construct three categories of variables: 1) illiquidity measures; 2) stock returns and related financial variables; and 3) measures of macroeconomic activity. The Internet Appendix provides additional details concerning data sources and the construction of certain measures.

2.1. Illiquidity measures

We construct a variety of firm-level measures of illiquidity and frictions at monthly and quarterly horizons for NYSE-listed stocks. Below we list and briefly describe these measures.

Roll measure (*ROLL*): The Roll (1984) liquidity measure is an estimate of the effective bid-ask spread based on the serial covariance of successive price movements.[7]

Corwin and Schultz measure (*CS*): The Corwin and Schultz (2012) liquidity measure uses high and low daily price quotes to estimate bid-ask spreads. The key idea behind the proxy is that, although the high-low ratio reflects both variance and the spread, the variance component is proportional to the return interval, while the spread component is not.

Fong, Holden, and Trzcinka measure (*FHT*): The Fong et al. (2017) liquidity measure incorporates two key elements of transaction costs: return volatility and the proportion of zero returns. The measure represents a simplification of a more general variation proposed by Lesmond et al. (1999).

Effective tick measure (*TICK*): The effective tick measure proposed by Holden (2009) and Goyenko et al. (2009) is based on observable price clustering.

Zeros measure (*ZEROS*): This liquidity measure, proposed by Lesmond et al. (1999), is based on the proportion of trading days with zero returns. A variation on this measure is defined as the fraction of zero return days among trading days with nonzero trading volume. Key results in the paper are robust to the version of zeros measure used.

Amihud illiquidity measure (*AMI*): Amihud (2002) proposes a measure of price impact based on the ratio of the daily absolute return to daily trading volume. The measure is computed as the average of this ratio over a certain time interval (e.g., month, quarter), including only days with positive trading volume. The measure captures the magnitude of daily price response per dollar of trading volume. To control for the effects of inflation on dollar trading volume, we construct *AMI* in real terms. However, we obtain similar results if we construct *AMI* in nominal terms. Following Brennan et al. (2013), we also construct a version of the Amihud measure based on the ratio of absolute daily returns to daily turnover, calculated as daily share volume divided by total shares outstanding (*AMITO*), in order to separate turnover from firm size effects.

Pastor and Stambaugh measure (*PS*): The Pástor and Stambaugh (2003) measure is a measure of price impact based on the extent of volume-related return reversal. The original measure proposed by Pástor and Stambaugh (2003) is one of liquidity (larger values indicate more liquidity). Our PS measure multiplies the original measure by negative one to convert to a measure of illiquidity.

Hou and Moskowitz measure (*HM*): The Hou and Moskowitz (2005) measure attempts to capture the delay with which stock prices respond to information. The measure is computed based on regressions of firm returns on current and lagged market returns. All else equal, the measure is larger when lagged market returns explain more variation in firm returns (relative to current market returns). Hou and Moskowitz (2005) interpret their measure as capturing a variety of market frictions including effects related to incomplete information, short-sale constraints, asymmetric information, taxes, and liquidity. Hou and Moskowitz (2005) find that the

[7] The Roll liquidity proxy is only defined when the first-order serial covariance of daily returns for the corresponding period is negative. Empirically, the first-order sample covariance can be positive. In such cases, we set the Roll proxy to zero. The Internet Appendix briefly discusses alternative conventions. Our main results are robust to the particular approach employed.

cross-sectional premium associated with their delay measure relates to measures of investor recognition, rather than traditional measures of liquidity.

Effective spread (*ESPD*): A common measure of the cost dimension of market liquidity using high-frequency (intradaily) trading data is the percent *effective spread*, defined as

$$ESPD = 2\,\frac{|P_k - M_k|}{M_k}, \tag{1}$$

where P_k represents the price of the kth trade, and M_k represents the midpoint of the best bid and offer (BBO) prevailing at the time of the kth trade. At the firm or asset level, this measure is aggregated over some unit of time (trading day, month, quarter) by computing the dollar-volume-weighted average of the percent effective spread for the time interval. The effective spread represents an estimate of the round trip cost of trading over the corresponding interval. We obtained this measure for all NYSE firms over the period 1993–2007 as a benchmark for evaluating the performance of the various low-frequency spread proxies.

For each firm-level measure of market illiquidity, we construct an aggregate measure by computing the equally weighted average across securities for each month or quarter. This procedure follows prior studies, including Jones (2002), Amihud (2002), Pástor and Stambaugh (2003), and Næs et al. (2011).

Most of the aggregate illiquidity measures are strictly positive. For these measures, it is convenient to apply a logarithmic transformation, which can be motivated by both statistical and economic considerations. From a statistical perspective, raw aggregate liquidity measures are positively skewed and leptokurtotic, and the logarithmic transformation dampens these deviations from the Gaussian distribution. From an economic perspective, we show in Section 3.2 that the logarithm of many illiquidity proxies can be decomposed into the sum of a component capturing aggregate volatility and a residual. Lowercase notation indicates log-transformed variables. For example, $roll_t \equiv \ln(ROLL_t)$.

2.2. Financial and macroeconomic variables

Our primary stock return measure (*ret*) is the log excess return on the NYSE index with the 3-month T-bill rate serving as the reference return. We also consider a number of financial variables often used to forecast stock returns. Among these, a crucial variable is aggregate volatility, since a central focus of our paper involves contrasting the volatility-related and non-volatility-related forecasting content in various illiquidity measures. Our primary measure of aggregate volatility is the cross-sectional average (for NYSE firms) of the natural logarithm of the sample standard deviation returns within the corresponding month or quarter, denoted *vol*. The motivation behind this particular measure of aggregate volatility will be made clear in Section 3.2 of the paper. We also construct several variables related to interest rates. The stochastically detrended short-term interest rate [Campbell (1991), denoted *TBL*] is the current 3-month T-bill rate less a (backward) two-year moving average. The term spread (*TERM*) is the difference between the 10-year Treasury yield and the yield on the 3-month Treasury bill. The default spread (*DEF*) is the yield on Moody's Baa-rated corporate bonds minus the yield on Moody's Aaa-rated corporate bonds. The commercial paper to Treasury spread (*CPSP*) is the spread between the 3-month commercial paper rate and the rate on 3-month Treasury bills. A third set of variables includes traditional financial ratios and a measure of aggregate equity issuance. *BMKT* is the ratio of book value to market value for the Dow Jones Industrial Average. The variables *ep* (*dp*) denote the difference of the log of earnings (dividends) and the log of prices for the Standard and Poor's (S&P) 500 index. Net equity issuance (*NTIS*) is the 12-month moving sum of net issuance by NYSE-listed stocks divided by the total end-of-year market capitalization of NYSE stocks. All financial variables are obtained over the period 1926–2015 at both the monthly and quarterly frequencies.

We measure output at the monthly frequency using real, seasonally adjusted industrial production (*IP*). These data cover the period 1926.1–2015.12, although we focus on the post-war period for comparability with other macroeconomic measures. As an alternative measure of business conditions, we examine the monthly, seasonally adjusted US civilian unemployment rate (*UE*), available over the period 1948.2–2015.12. At the quarterly frequency, we examine output, unemployment, investment, and consumption, following Næs et al. (2011). The output measure is real, seasonally adjusted gross domestic product (*GDP*). Unemployment (*UE*) is the seasonally adjusted unemployment rate. Our measure of investment (*INV*) is real, gross private domestic investment (seasonally adjusted). Consumption (*CONS*) is real, seasonally adjusted personal consumption expenditures. These data are available over the period 1947.1–2015.4, with the exception of *UE* which is available beginning in 1948. The macroeconomic variables either trend over time, or exhibit considerable persistence and potential structural shifts. For this reason, we focus on growth rates. For example, the variable Δip denotes the growth rate of industrial production, computed as log-differences.

3. Aggregate illiquidity

This section analyzes the dynamics of aggregate illiquidity. Sections 3.1 characterizes the time-series behavior of the various aggregate illiquidity measures. Section 3.2 discusses approaches to decomposing illiquidity measures into a component related to aggregate volatility and a residual. Section 3.3 discusses structural shifts in many illiquidity measures around the turn of the century related to tick-size reductions.

3.1. Time-series properties of aggregate illiquidity measures

Many of the low-frequency illiquidity measures aim to measure the round trip cost of trading (bid-ask spread). Fig. 1 plots time series for several alternative spread proxies. The aggregate time series comove closely for most of

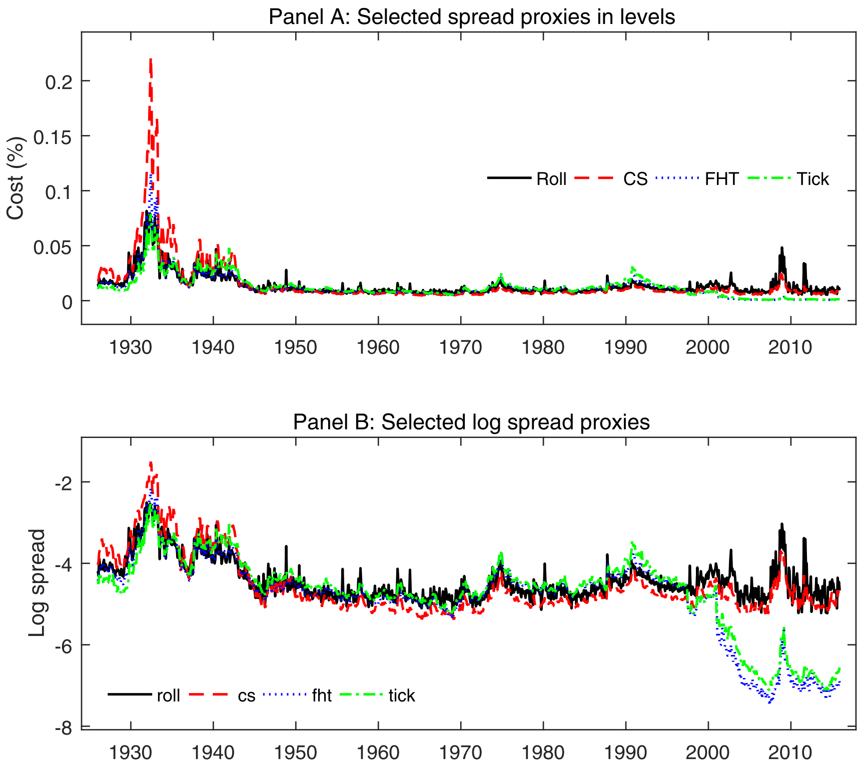

Fig. 1. Alternative effective spread proxies: 1926–2015. This figure shows monthly time series for selected low-frequency proxies for the aggregate effective spread. The top panel shows the spread proxies in levels, while the bottom panel plots log versions of the series. *ROLL* is the proxy of Roll (1984). *CS* is the proxy proposed by Corwin and Schultz (2012). *FHT* is the proxy proposed by Fong et al. (2017). *TICK* is the effective tick measure proposed by Holden (2009). Log versions of series are denoted using lowercase letters (e.g., *roll*). We show results for selected measures to enhance readability. The Internet Appendix provides time-series plots for other proxies.

the sample period. The affinity among the various spread proxies is particularly high during a 60-year period stretching from the 1940s through the 1990s. Consistent with Jones (2002), the aggregate spread series exhibit cyclical behavior. Spreads tend to increase during economic recessions and financial crises. The aggregate spread increases dramatically around the Great Depression, although the extent of the increase differs somewhat across alternative proxies.

The spread proxies diverge in recent years. The *FHT* and *TICK* measures appear to shift downward in the late 1990s and again in the early 2000s. These shifts are particularly apparent for the log-transformed versions of the series (Panel B of Fig. 1). Interestingly, the *ROLL* and *CS* proxies do not exhibit a clear downward break at this time. The Internet Appendix provides a time-series plot comparing the aggregate effective spread computed using intraday Trade and Quote (TAQ) data to the various low-frequency spread proxies for the sub-period 1993–2007. The effective spread computed using TAQ data conforms closely to the *FHT* and *TICK* proxies throughout this period. In contrast, *ROLL* and *CS* are upward-biased measures of the aggregate effective spread from the late 1990s onward.

The Internet Appendix provides time-series plots for various other illiquidity proxies we consider. Most illiquidity proxies fluctuate inversely with the business cycle, and tend to rise during recessions and financial crises. Notably, the zeros proxy exhibits two sharp downward structural breaks, one following June 1997 and the other following January 2001, coinciding with minimum tick-size reductions for the NYSE. The stark pattern for the zeros proxy lends credence to the hypothesis that downward structural shifts for other illiquidity measures around the turn of the century relate to tick-size reductions.

Table 1 presents descriptive statistics for illiquidity measures and other selected variables at the monthly frequency over the sample period 1948.2–1997.6.[8] Most illiquidity measures are positively skewed and leptokurtotic, even after applying a logarithmic transformation. They are also persistent. First-order sample autocorrelations range from 0.33 (the Hou–Moskowitz measure) to 0.98 (the effective tick measure). Autocorrelations for the illiquidity measures die out slowly. The table shows the sample autocorrelation at lag 36 (3 years), along with the corresponding implied autocorrelation at lag 36 under an AR(1) process. The sample autocorrelation at a 36-month lag is often much larger than that implied by an AR(1) process. This suggests that aggregate illiquidity measures exhibit *long memory* properties. The autocovariances of a long memory process decay at a slower rate relative to the geometric de-

[8] This period is chosen to minimize the impact of apparent structural shifts occurring outside of this period on correlation estimates and measures of persistence.

Table 1
Descriptive statistics: raw illiquidity measures and selected macroeconomic and financial variables.

This table presents summary statistics for aggregate illiquidity measures and selected macroeconomic and financial variables. Panel A presents the sample mean, standard deviation, skewness, and kurtosis for each variable, along with several statistics characterizing the degree of persistence in the time series. Sample means and standard deviations for the variables Δip, Δue, and *ret* are monthly percentages. $\hat{\rho}_1$, and $\hat{\rho}_{36}$ are the sample autocorrelations at lags 1 and 36, respectively. $\hat{\rho}_1^{36}$ is the correlation at lag 36 implied by the sample autocorrelation AC(1), assuming a first-order autoregressive (AR(1)) process for the corresponding variable. $\hat{d}$ is the exact local Whittle estimate of the long memory parameter d proposed by Shimotsu and Phillips (2005). Asymptotic standard errors are provided for this statistic. The top portion of Panel B provides selected pairwise sample correlation estimates. The bottom portion of Panel B shows the estimated correlation between shocks to the corresponding variable and shocks to either excess returns or growth in industrial production. See Section 2 of the paper or the Internet Appendix for variable definitions. The sample period is 1948.2–1997.6, selected to avoid reporting descriptive statistics over periods of structural change related to tick-size reductions.

	Illiquidity measures									Other variables			
	roll	*cs*	*fht*	*tick*	*zeros*	*ami*	*amito*	*PS*	*hm*	*ret*	Δip	Δue	*vol*
Panel A: Univariate descriptive statistics													
Mean	−4.66	−4.91	−4.63	−4.54	−1.67	0.68	3.27	0.01	−0.45	0.58	0.29	0.07	−3.93
St. dev.	0.23	0.24	0.28	0.30	0.17	0.71	0.32	0.02	0.16	4.06	1.05	4.19	0.21
Skew.	0.78	0.62	0.34	0.61	−0.51	0.14	0.55	2.65	−1.52	−0.65	0.30	0.46	0.70
Kurt.	4.08	3.10	3.06	3.64	3.22	3.55	3.89	17.77	6.75	5.89	7.49	6.96	4.89
$\hat{\rho}_1$	0.66	0.96	0.97	0.98	0.88	0.96	0.88	0.37	0.33	0.05	0.40	0.14	0.83
$\hat{\rho}_{36}$	0.22	0.45	0.54	0.48	0.34	0.38	0.12	0.03	0.07	0.01	−0.02	−0.09	0.32
$\hat{\rho}_1^{36}$	0.00	0.21	0.30	0.41	0.01	0.26	0.01	0.00	0.00	0.00	0.00	0.00	0.00
$\hat{d}$	0.66	0.83	0.87	0.89	0.76	0.85	0.63	0.35	0.36	−0.03	0.02	0.12	0.66
SE($\hat{d}$)	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
Panel B: Selected correlations													
roll	1.00	0.80	0.76	0.77	0.25	0.70	0.40	0.40	−0.34	−0.12	−0.21	0.16	0.61
ami	0.70	0.87	0.83	0.90	0.66	1.00	0.50	0.15	−0.21	−0.01	−0.21	0.19	0.31
ret	−0.12	−0.02	−0.01	−0.02	0.10	−0.01	−0.21	−0.32	0.39	1.00	0.00	−0.03	−0.14
Δip	−0.21	−0.23	−0.19	−0.20	−0.10	−0.21	−0.24	−0.12	0.07	0.00	1.00	−0.51	−0.16
Δue	0.16	0.20	0.15	0.16	0.09	0.19	0.25	0.13	−0.09	−0.03	−0.51	1.00	0.12
vol	0.61	0.55	0.64	0.57	−0.30	0.31	0.31	0.46	−0.37	−0.14	−0.16	0.12	1.00
$\hat{\rho}(\epsilon_{ret}, \epsilon_X)$	−0.27	−0.37	−0.33	−0.48	0.04	−0.41	−0.46	−0.33	0.48				
$\hat{\rho}(\epsilon_{\Delta ip}, \epsilon_X)$	−0.07	−0.03	−0.07	−0.05	−0.09	−0.09	−0.08	−0.03	−0.01				

cay rate associated with autocovariances for short memory processes, such as stationary autoregressive moving average (ARMA) processes.[9] Table 1 reports estimates of the long memory parameter d using the local Whittle estimator (Shimotsu and Phillips, 2005). The d estimates are statistically greater than zero for all illiquidity measures, and most estimates fall in the range 0.3–0.9. Consistent with numerous previous studies, estimates in Table 1 imply that volatility is also a long memory process.[10] In contrast, output growth, unemployment growth, and stock market returns yield d estimates that are close to zero.

Panel B of Table 1 displays sample correlations among the illiquidity proxies. Alternative spread proxies tend to be highly correlated over the 1948–1997 period. Most other illiquidity measures, including Amihud measures and the *zeros* measure, are positively correlated with the spread measures and with each other. Interestingly, the *hm* price delay measure is *negatively* correlated with spread and price impact measures, suggesting that this measure emphasizes other sources of frictions.

The illiquidity proxies are generally negatively correlated with the growth rate in industrial production and positively correlated with unemployment growth, as intuition suggests. Stock market returns are negatively correlated with most illiquidity measures: a positive market return tends to coincide with an improvement in aggregate liquidity. Stock market volatility is positively correlated with most illiquidity measures—a relation we explore below in further detail. The price delay proxy *hm* correlates positively with contemporaneous stock returns, and negatively with stock market volatility. Hou and Moskowitz (2005) find that stocks with large price delay effects tend to be small, with limited institutional ownership and analyst coverage. To the extent that the aggregate *hm* measure captures attentional effects, it is plausible that average investor attention increases in volatile, turbulent market conditions associated with recessions.

The last two rows of Table 1 report estimates of the contemporaneous correlation between *shocks* to illiquidity measures and shocks to either excess stock returns or

[9] The autocorrelations associated with a long memory process die out at a hyperbolic rate, in contrast to the faster exponential rate for short memory processes. Granger and Joyeux (1980) and Hosking (1981) show that a stationary long memory time series can be modeled by extending the standard autoregressive integrated moving average (ARIMA) class of models to permit *fractional integration*. Such a process may be written as

$$(1-L)^d(x_t-\mu_x)=w_t,$$

where $\mu_x = E(x_t)$, w_t is a short memory, zero mean time series, $-0.5 < d < 0.5$ and (setting $\mu_x = 0$), $(1-L)^d x_t$ has the interpretation $1 - dx_{t-1} + \frac{d(d-1)}{2!}x_{t-2} - \cdots$, with d the so-called fractional differencing parameter. The case $d = 0$ corresponds to the absence of long memory (a short memory series). Values of d above 0.5 are also possible; however, in such cases the resulting process is no longer covariance stationary. The case $d = 1$ corresponds to an integrated process for x_t.

[10] Examples of studies documenting long memory in stock market volatility include Ding et al. (1993), Baillie et al. (1996), and Andersen and Bollerslev (1997).

growth in industrial production. To obtain these estimates, we first apply a univariate time-series filter to each variable, and then compute the sample correlation for the resulting residuals. Shocks for illiquidity measures are based on an autoregressive fractionally differenced moving average (ARFIMA) model of order (1, d, 0). This model includes both the long memory parameter d and a short memory component modeled as an AR(1) process. Model parameters are estimated using the conditional sum of squares (CSS) approach (Beran, 1995), discussed further in the Internet Appendix. Since there is little persistence in stock returns, return shocks are simply deviations from the mean. Shocks to growth in industrial production are based on an autoregressive model with four lags included.

Shocks to most illiquidity measures are negatively correlated with contemporaneous shocks to excess stock returns. This is consistent with theoretical predictions (see, e.g., Acharya and Pedersen, 2005). The sign of the correlation estimate is positive rather than negative for the *hm* measure. The magnitude of most correlation estimates falls in the range 0.25–0.5. There appears to be little correlation between shocks to illiquidity measures and shocks to growth in industrial production. We reference these correlation estimates later in discussing the econometric properties of predictive regressions involving illiquidity measures.

3.2. Extracting a volatility component from illiquidity measures

Most measures of transaction costs are positively correlated with stock market volatility (see Table 1). Below we show that several standard illiquidity measures are *mechanically* related to volatility, and that the associated relation is log-linear. This provides a convenient basis for decomposing aggregate illiquidity measures into a volatility component and a residual. The decomposition we make is empirical in nature. However, motivation for our decomposition also stems from theory. For example, the price impact parameter λ in the seminal model of Kyle (1985) is log-linear in the volatility of fundamental asset value.

As a first example, consider the spread proxy proposed by Fong et al. (2017). For the ith firm in period t, the measure is defined as:

$$FHT_{i,t} = 2\hat{\sigma}_{i,t}\Phi^{-1}\left(\frac{1+ZEROS_{i,t}}{2}\right), \tag{2}$$

where $\hat{\sigma}_{i,t}$ is the standard deviation of the ith firm's daily returns over time interval t (e.g., a month or quarter), $ZEROS_{i,\,t}$ is the fraction of zero return days out of the total trading days during time interval t, and $\Phi^{-1}(\cdot)$ is the inverse cumulative normal distribution. Assuming $ZEROS_{i,\,t} > 0$ and taking logarithms yields:

$$fht_{i,t} = \ln(2) + \ln(\hat{\sigma}_{i,t}) + g(ZEROS_{i,t}), \tag{3}$$

where $g(ZEROS_{i,t}) \equiv \ln\left(\Phi^{-1}\left(\frac{1+ZEROS_{i,t}}{2}\right)\right)$. Averaging across firms gives

$$\overline{fht_t} = \ln(2) + vol_t + \overline{g(ZEROS_{i,t})}, \tag{4}$$

where overbars indicate cross-sectional averages and $vol_t \equiv \overline{\ln(\hat{\sigma}_{i,t})}$, the cross-sectional average of the logarithm of firm volatility in period t. This shows that, ignoring the constant $\ln(2)$, the aggregate quantity $\overline{fht_t}$ is the sum of two components. The first is the cross-sectional average of the log of estimated volatility at the firm level (vol_t), and the second is the cross-sectional average of a nonlinear transformation of the *ZEROS* illiquidity measure.

The variable vol_t is constructed from firm-level return standard deviations that incorporate idiosyncratic as well as systematic risk. However, the *aggregate* measure vol_t produced by taking cross-sectional averages of these firm-level quantities is closely related to traditional measures of stock market volatility. The top panel of Fig. 2 provides a visual illustration of this point. The figure compares a time-series plot of vol_t with a more traditional measure of stock market volatility: the logarithm of the monthly standard deviation of daily returns on the S&P 500 Index (French et al., 1987). Both variables are standardized prior to plotting to facilitate comparison. The series comove very closely, illustrating that fht_t is mechanically related to a quantity that is, essentially, a measure of market volatility.

Other illiquidity measures also mechanically relate to vol_t. For example, the Roll (1984) proxy takes the form:

$$ROLL_{i,t} = 2\sqrt{-\widehat{\mathrm{Cov}}(R_{d,i,t}, R_{d-1,i,t})}, \tag{5}$$

where $R_{d,\,i,\,t}$ ($R_{d-1,i,t}$) is the return for the ith firm on trading day d (day $d-1$) of time interval t and $\widehat{\mathrm{Cov}}(\cdot)$ represents the sample covariance. Taking logs and averaging across firms provides a decomposition similar to Eq. (4):

$$\overline{roll_t} = \ln(2) + vol_t + (1/2)\overline{\ln\left(-\hat{\rho}_1(R_{t,i})\right)}, \tag{6}$$

where $\hat{\rho}_1$ denotes the sample autocorrelation at lag one, and it is implicitly assumed that $\hat{\rho}_1(R_{t,i}) < 0$. In the Internet Appendix, we show that, under certain assumptions, it is also possible to analytically derive a decomposition similar to Eqs. (4) and (6) for Amihud-type measures.

We argue that, in many applications, it is of interest to separate the market volatility component from aggregate illiquidity measures. Consider, for example, the relation between illiquidity and the equity premium. Standard dynamic asset pricing models, such as the intertemporal capital asset pricing model (ICAPM) of Merton (1973), imply a positive equilibrium relationship between the conditional market risk premium and conditional volatility even in the absence of transactions costs. The volatility component embedded in conventional aggregate illiquidity measures also potentially complicates inference concerning connections between illiquidity and economic activity, because previous literature finds that market volatility forecasts future economic activity (see, e.g., Campbell et al., 2001; Fornari and Mele, 2013). Finally, in a cross-sectional asset pricing context, Bansal et al. (2014) emphasize the role of market volatility as a separate, priced risk factor. Because standard aggregate illiquidity measures contain a volatility component, the apparent pricing of illiquidity risk could solely be attributable to volatility risk, which can be priced in the absence of market frictions (Bansal et al., 2014).

The explicit decomposition results above motivate the following simple, generalized decomposition for an arbitrary illiquidity measure ILQ_t:

$$\widetilde{ILQ}_t = ILQ_t - vol_t. \tag{7}$$

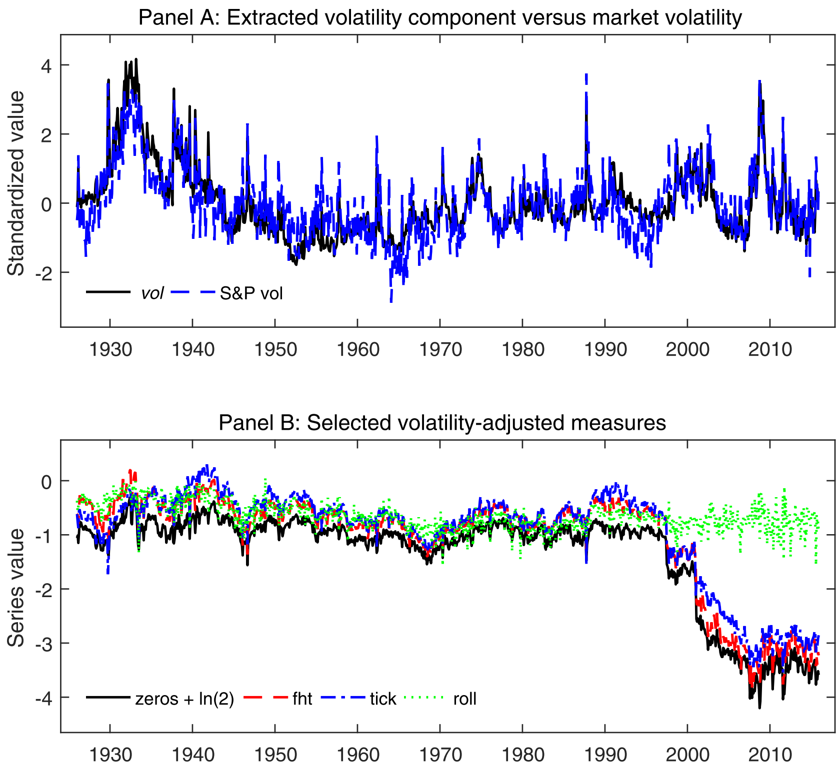

Fig. 2. Decomposing illiquidity measures into a volatility component and a residual. This figure illustrates the effects of decomposing selected aggregate illiquidity measures into a volatility component and a residual. Panel A compares the time series *vol* with a more conventional measure of stock market volatility. Specifically, the series denoted 'S&P vol' is the logarithm of the monthly volatility of the S&P 500 Index, where the monthly volatility is constructed as the square root of the sum of squared daily S&P 500 returns. Panel B depicts monthly time series for various (log) spread proxies adjusted for aggregate volatility. As a reference point for comparison, the figure includes the raw illiquidity measure $zeros + \ln(2)$, where the addition of $\ln(2)$ is based on Eq. (4). The sample period is 1926–2015.

This is useful for proxies such as *tick* and *cs*, for which it is more difficult to explicitly derive log-linear decompositions.[11] Since these proxies also measure the effective spread, one can conjecture that they are also log-linear in volatility and one may simply subtract log volatility to obtain an adjusted measure. Two remarks are in order. First, nothing in the derivation of decompositions of the form of Eq. (7) implies that the associated decomposition is orthogonal, in the sense that vol_t and the residual $\widetilde{ILQ}_t$ are uncorrelated. Second, it is not clear ex ante that a decomposition of the form of Eq. (7) is sensible for all illiquidity measures we consider. In fact, the decomposition of Eq. (4) suggests that the *zeros* measure *itself* can be interpreted as a volatility-adjusted measure of the aggregate effective spread, and subtracting vol_t for this measure would unnecessarily introduce a volatility component.

An alternative, data-driven decomposition of an arbitrary illiquidity measure ILQ_t can be based on the linear regression:

$$ILQ_t = \alpha + \beta\, vol_t + \epsilon_t, \tag{8}$$

where $\beta \equiv \mathrm{Cov}(ILQ_t, vol_t)/\mathrm{Var}(vol_t)$. The regression residual ϵ_t in Eq. (8) is uncorrelated with vol_t by virtue of the definition of β. Let $\hat{\alpha}$ and $\hat{\beta}$ denote estimates of the parameters α and β in Eq. (8). We define a volatility-adjusted illiquidity measure as the residual: $ILQ_t^{\perp} \equiv \hat{\epsilon}_t$, where $\hat{\epsilon} = ILQ_t - \hat{\alpha} - \hat{\beta}\, vol_t$. The notation emphasizes that this form of volatility-adjustment orthogonalizes the illiquidity measure with respect to aggregate log volatility. In the special case that $\beta = 1$ in Eq. (8), the analytical decomposition of Eq. (7) becomes an *orthogonal* decomposition of the spread, and is equivalent to the linear projection approach up to an intercept, in a population sense. Consequently, it is of interest to test the restriction $\beta = 1$ in the data.

Table 2 shows results for analytical and projection-based decompositions of illiquidity measures. For each measure, the left-hand columns of Table 2 report statistics for regressions of ILQ_t on vol_t. Specifically, the table shows $\hat{\beta}$, the ordinary least squares (OLS) estimate of β, as well as the associated standard error and t-statistics for tests of

[11] The volatility adjustment of Eq. (7) subtracts vol_t from an aggregate illiquidity measure, such as fht_t. For this reason, there is a subtle difference between, e.g., $\widetilde{fht}_t$ and the quantity $\overline{fht}_t - vol_t$ based on Eq. (4). Constructing $\widetilde{fht}_t$ rather than $\overline{fht}_t - vol_t$ resolves a technical issue that arises due to the fact that the quantity $g(ZEROS_{i,t})$ in Eq. (4) is undefined when $ZEROS_{i,t} = 0$. This case occurs in the data, particularly in recent years following decimalization. The two quantities differ by a Jensen's inequality correction term, and the substitution ensures that $\widetilde{fht}$ is well-defined. Similar remarks pertain to the Roll measure. When $\hat{\rho}_1(R_{t,i}) > 0$ in Eq. (6), the second term on the right-hand side of Eq. (6) is undefined. Constructing $\widetilde{roll}$ effectively substitutes $roll_t = \ln(\overline{ROLL}_{i,t})$ for $\overline{roll}_t$ in Eq. (6).

Table 2
Extracting a volatility component from illiquidity measures.

The table presents results for projection-based and analytical decompositions of illiquidity measures into a volatility component and a residual. ILQ_t denotes a generic illiquidity measure. Projection-based decompositions are based on regressions of the form

$$ILQ_t = \alpha + \beta vol_t + \epsilon_{t+1},$$

where vol_t denotes (log) average volatility. For each illiquidity measure, the table reports $\hat{\beta}$, the OLS estimate of β, as well as the associated standard error. t-statistics are provided for the typical null of $\beta = 0$ and for the null $\beta = 1$, which is imposed under the analytical decomposition. The fifth column of the table shows the R^2-value associated with the OLS regression, expressed as a percentage. $\hat{\beta}^{NB}$ denotes the narrow-band least squares estimate of β with a bandwidth of $T^{0.6}$. $\hat{\rho}(ILQ^{\perp}, \widetilde{ILQ})$ denotes the sample correlation between the volatility-adjusted proxy $ILQ_t^{\perp}$, defined as the residuals from the regression of ILQ_t on vol_t, and the analytical volatility-adjusted proxy $\widetilde{ILQ}_t$, defined as $ILQ_t - vol_t$. $R^2_{\widetilde{ILQ}}$ denotes the R^2-value associated with the analytical decomposition, i.e., the quantity $1 - \frac{\text{Var}(\widetilde{ILQ}_t)}{\text{Var}(ILQ_t)}$, expressed as a percentage. This value need not be positive, because the analytical adjustment may produce a fit inferior to the sample average. The corresponding values for *PS* and *hm* are in fact negative and large (in magnitude). To conserve space, we indicate this as "$\ll 0$." The sample period is 1926–2015. See Section 2 of the paper or the Internet Appendix for variable definitions.

	OLS estimation of β							
ILQ	$\hat{\beta}$	*SE*	$t(\beta = 0)$	$t(\beta = 1)$	R^2	$\hat{\beta}^{NB}$	$\hat{\rho}(ILQ^{\perp}, \widetilde{ILQ})$	$R^2_{\widetilde{ILQ}}$
roll	1.19	0.07	18.11	2.95	75.55	1.26	0.96	73.55
cs	1.62	0.17	9.74	3.72	64.65	1.80	0.89	55.20
fht	1.10	0.30	3.63	0.33	16.57	1.18	1.00	16.43
tick	0.95	0.26	3.57	−0.21	14.52	1.01	1.00	14.48
zeros	−0.05	0.30	−0.16	−3.49	0.04	0.01	0.92	−18.95
ami	1.43	0.23	6.15	1.86	30.19	1.55	0.98	27.43
amito	0.71	0.19	3.78	−1.56	17.80	0.76	0.98	14.76
PS	0.02	0.01	3.17	−154.80	7.31	0.02	0.07	$\ll 0$
hm	−0.26	0.07	−3.57	−17.10	14.78	−0.19	0.45	$\ll 0$

the null of $\beta = 0$ and the alternative null of $\beta = 1$. The null $\beta = 0$ reflects a situation in which the corresponding illiquidity measure is uncorrelated with aggregate (log) volatility. The alternative null $\beta = 1$ corresponds to the value imposed by the analytical decomposition formula. The fifth column of the table shows the R^2-value associated with the OLS regression, expressed as a percentage.

OLS estimates of $\hat{\beta}$ are positive and statistically different from zero for most proxies. One exception is the *zeros* measure, for which the null that $\beta = 0$ cannot be rejected. Recall that the *zeros* measure is closely related to a volatility-adjusted version of the spread proxy *fht* (see Eq. (4)). The fact that the null of $\beta = 0$ cannot be rejected for *zeros* indicates that $\widetilde{fht}$ is essentially uncorrelated with aggregate volatility, i.e., the components *vol* and $\widetilde{fht}$ represent a roughly orthogonal decomposition of the *fht* illiquidity measure.[12] A second exception is the *hm* measure, for which $\hat{\beta}$ is negative and significant. This is not surprising in light of previous evidence that *hm* is negatively correlated with traditional illiquidity measures. Although volatility is positively correlated with most illiquidity measures, there is heterogeneity across measures with regard to the importance of the volatility component, as measured by the regression R^2-value. For example, aggregate volatility explains a majority of the variation in the *roll* and *cs* measures: the regression R^2-values for these measures are around 75% and 65%, respectively. In contrast, the R^2-values associated with the alternative spread proxies *fht* and *tick* are roughly 15%. OLS estimates of β are close to one for the spread and price impact measures, particularly *fht* and *tick*. Indeed, the null hypothesis that $\beta = 1$ cannot be rejected for these measures. The β estimate for the *amito* price-impact measure is also statistically indistinguishable from one. Estimates of β are larger than one for *roll*, *cs*, and *ami*, and the null that $\beta = 1$ is rejected.

The apparent long memory behavior of both volatility and aggregate illiquidity raises potential concerns regarding the appropriateness of OLS as an estimator of β. As a robustness check, we also estimate β using a narrow-band least squares estimator, denoted $\hat{\beta}^{NB}$, proposed by Robinson (1994). This is a frequency-domain least squares estimator such that estimation is carried out over a degenerating band of frequencies around the origin. Robinson and Marinucci (2003) and Christensen and Nielsen (2006) provide asymptotic distributional results. To implement the estimator, it is necessary to choose a bandwidth parameter m, which we set to the integer part of $T^{0.6}$, where T is the sample size. The narrow-band least squares estimates of β are often slightly larger than the OLS estimates; however, differences are not pronounced.

The final two columns of Table 2 show that, for most proxies, the analytical and projection-based approaches yield very similar volatility-adjusted measures. The statistic $\hat{\rho}(ILQ^{\perp}, \widetilde{ILQ})$ denotes the sample correlation between the volatility-adjusted proxy $ILQ_t^{\perp}$ and the analytical volatility-adjusted proxy $\widetilde{ILQ}_t$. With the exception of *PS* and *hm*, the

[12] This finding is consistent with Bekaert et al. (2007), who examine the relation between the zeros illiquidity measure and volatility in a set of emerging markets. They find the correlation between the two to be small in magnitude on average across countries and conclude that the zeros measure captures "unique aspects of liquidity not entirely driven by the presence or absence of news in a particular period" (p. 1791).

correlation between the two alternative volatility-adjusted measures is close to one. $R^2_{\widetilde{ILQ}}$ represents the implicit R^2-value associated with the analytical decomposition, i.e., the quantity $1 - \frac{\text{Var}(\widetilde{ILQ}_t)}{\text{Var}(ILQ_t)}$, expressed as a percentage. This value need not be positive, because it is possible that the analytical adjustment produces a fit inferior to the sample average. Empirically, the R^2-value associated with the simple analytical decomposition is close to (but necessarily smaller than) the regression R^2-value for most measures. Not surprisingly, the analytical decomposition does not work well for the measures *PS* and *hm*. The β estimate for *PS* is close to zero rather than one, while $\hat{\beta}$ for *hm* is negative. The null hypothesis that $\beta = 1$ is strongly rejected for both measures.

The bottom panel of Fig. 2 shows selected volatility-adjusted spread proxies over the period 1926–2015. The volatility-adjusted measures are based on the analytical decomposition of Eq. (7), but results are similar for measures based on the regression approach. For comparison, the plot includes the *unadjusted* measure $zeros + \ln(2)$. The series $\widetilde{fht}$ is extremely similar to the quantity $zeros + \ln(2)$, consistent with the analytical decomposition of Eq. (4). Despite the lack of an explicit algebraic decomposition, $\widetilde{tick}$ is also very similar to $zeros + \ln(2)$. Both $\widetilde{fht}$ and $\widetilde{tick}$ exhibit downward shifts around tick-size reductions in 1997 and 2001, similar to the *zeros* measure. This implies that the downward structural breaks associated with *fht* and *tick* are driven by the residual component after adjusting for volatility. The series $\widetilde{roll}$ also comoves closely with $zeros + \ln(2)$ until the late 1990s; however, it does not shift downward at this time.

Subsequent empirical analyses consider volatility-adjusted illiquidity measures. To minimize potential look-ahead bias concerns, we primarily rely on volatility-adjusted measures using the simple analytical approach that subtracts vol_t. Two exceptions are the *PS* and *hm* proxies, which we adjust using the OLS regression approach based on Eq. (8). Results in Table 2 indicate that the projection-based decomposition is more appropriate for these proxies. Finally, because the *zeros* measure appears to be already orthogonal to volatility, we make no adjustment for this variable. We obtain similar results in our main empirical tests for variations on the volatility-adjustment approach. For example, basing all volatility-adjusted quantities on the projection approach (with either OLS or narrow-band least squares estimates) yields similar results.

3.3. Break-adjusted illiquidity measures

Several illiquidity measures appear to exhibit structural shifts around tick-size reductions for the NYSE in 1997 and 2001. Effects are particularly pronounced for the *zeros* measure and several spread proxies. It is potentially prudent to adjust illiquidity measures for structural shifts around these tick-size reductions. There exists a long-running debate in the literature regarding the impact of tick-size changes on market liquidity. Tick-size reductions potentially benefit liquidity demanders when competition between liquidity providers leads to reductions in the bid-ask spread. But such reductions may also decrease the provision of liquidity (Harris, 1997). Harris (1994) and Seppi (1997) provide theoretical arguments suggesting that tick-size reductions can reduce, rather than improve, liquidity. A large body of empirical research indicates that tick-size reductions reduce quoted and effective spreads. However, there is also evidence that tick-size reductions reduce depth and potentially increase transactions costs on net. For example, together with the reduction in bid-ask spreads, Goldstein and Kavajecz (2000) find the quantity of shares posted as limit orders on the NYSE declined following the 1997 tick-size reduction. Jones and Lipson (2001) find that total trading costs for institutional investors who execute large trades actually increased after the tick-size reduction in 1997. Similar evidence has been documented for the Toronto Stock Exchange following its tick-size reduction in 1996 (see, e.g., Bacidore, 1997; Ahn et al., 1998). Finally, the Security and Exchange Commission's (SEC) recent implementation of a pilot program to experimentally *increase* the minimum tick-size for certain stocks reflects concerns that small tick sizes potentially hinder liquidity provision, particularly for stocks with small market capitalization. (See, e.g., https://www.sec.gov/news/pressrelease/2015-82.html.)

The illiquidity measures we construct represent proxies for imperfectly observable true liquidity conditions. It is certainly possible that tick-size reductions in 1997 and 2001 generated improvements in aggregate liquidity. However, the magnitude of reductions occurring over short time intervals for certain proxies seem implausibly large to correspond to 'true' liquidity improvements. As a concrete example, the *zeros* measures falls by roughly 35% over two successive months around decimalization in early 2001. While not discounting the possibility that latent aggregate liquidity improved over this period, any corresponding improvement is presumably considerably smaller in magnitude.

We consider several potential approaches to adjusting for structural shifts in illiquidity measures. A first approach adjusts for hypothesized level shifts in July 1997 and February 2001 driven by associated tick-size reductions.[13] Specifically, we estimate the magnitude of level shifts occurring at these points on an ex post basis using the full sample of data for each illiquidity measure. Break-adjusted series are constructed by subtracting these estimated break magnitudes from the illiquidity measure over the relevant post-break intervals. We refer to this procedure as the full sample (FS) break-adjustment method. A second variation on this approach proceeds in 'real time,' (RT) in the sense that we recursively test for and identify breaks using only data available at the time. Specifically, we recursively test for the presence of a level-shift occurring in the month corresponding to each tick-size reduction. The series is then adjusted based on the recursively estimated break magnitudes only if the null hypothesis of no break is rejected at the 5% level. See the Internet

[13] The specific implementation dates for the tick-size reductions were June 24, 1997 and January 29, 2001, respectively. Since these dates occur near the end of the corresponding month, we identify the subsequent months as the conjectured break points.

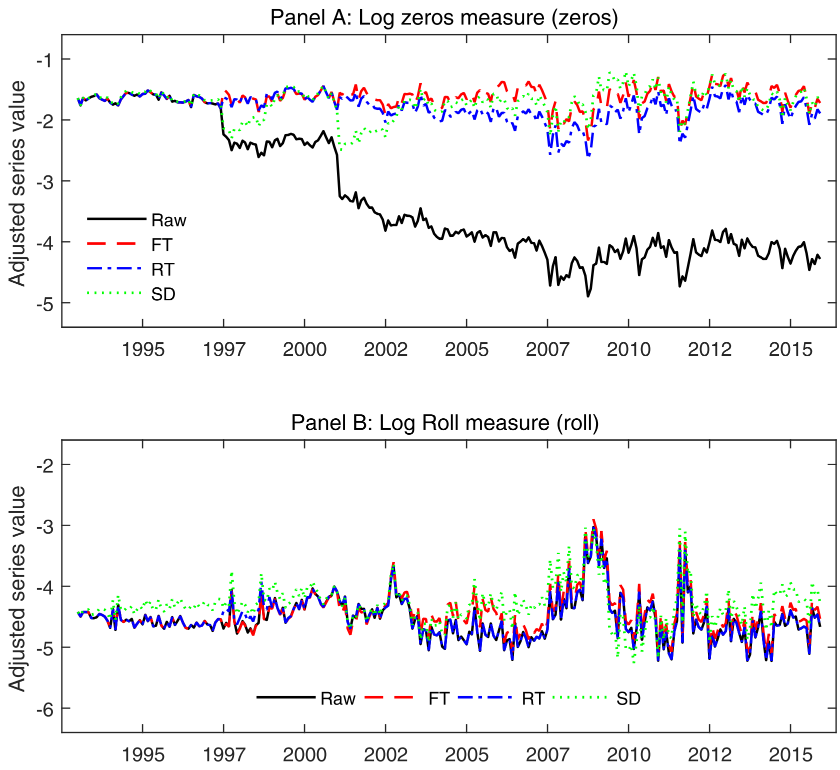

Fig. 3. Illustration of break-adjustment for selected measures. This figure compares raw illiquidity measures with break-adjusted series for the period 1993.1–2015.12. Panel A shows results for the *zeros* illiquidity proxy and Panel B shows results for the *roll* illiquidity proxy. For each illiquidity measure, the figure shows the unadjusted series ('Raw,' solid line), the adjusted series based on the full-sample approach ('FS,' dashed line), the adjusted series based on the real time approach involving recursive tests for level shifts at tick-size reduction points ('RT,' dash-dot line), and the adjusted series based on a stochastic detrending procedure using a two-year moving average ('SD,' dotted line). To facilitate comparison in the plot, the FS and SD adjusted series are level-adjusted so that their initial value in 1993.1 equals that of the raw series.

Appendix for additional details. The adjustment procedure implicitly assumes that the forecaster focuses on tick-size reduction points as potential break dates, and does so in real time. This assumption is plausible in our context, since the hypothesized break dates coincide with the onset of tick-size reductions, which were known well in advance of implementation.[14]

As a final, alternative approach to adjusting illiquidity measures for structure change, we compute stochastically detrended (SD) series. Stochastically detrended series are constructed as the difference between the current value of the series and a backwards moving average. It is relatively common to stochastically detrend the short-term Treasury bill rate in predictive regressions for stock returns (see, e.g., Campbell, 1991; Hodrick, 1992). The stochastic detrending procedure possesses several advantages: it is backward-looking and in this sense inherently real time, and it adjusts for broader types of nonstationary behavior.[15] We use a two-year window (24 months) for computing the backward moving average for the stochastically detrended series.

Fig. 3 illustrates the effects of applying break-adjustment procedures for two benchmark illiquidity measures: the *zeros* measure and the *roll* measure. The figure contrasts the FS, RT, and SD break-adjustment approaches. The top panel of Fig. 3 illustrates break-adjusted series for the *zeros* measure, which exhibits sharp, sudden decreases around tick-size reduction points. There are only modest differences between the FS and RT adjusted series, due primarily to the fact that the RT adjustment procedure detects breaks corresponding to tick-size reductions in 1997 and 2001 almost immediately. The SD adjustment gradually reacts to each break, in contrast to the rapid adjustment under the FS and RT approaches. The bottom panel of Fig. 3 illustrates the effects of adjustment for the *roll* measure, for which it is less clear whether structural shifts occur. Overall, differences between the raw series and the FS and RT adjusted series are modest. The RT series follows the raw series closely because, except for a brief pe-

[14] We also consider a variation of the real time procedure that imposes breaks at tick-size reduction points (without testing), and recursively adjusts series based on estimated break magnitudes. This procedure produces series that are similar to the approach involving testing. See the Internet Appendix for further discussion and selected results illustrating robustness.

[15] Campbell (1991) points out that a stochastically detrended series can be written as a triangular moving average of changes in the series. This implies that the detrended series is stationary in levels if the underlying series is stationary in differences. A potential disadvantage of the stochastic detrending procedure is that the adjusted series adapts only gradually to a sudden, permanent level shift.

riod shortly after the 1997 tick-size reduction, the break tests fail to reject the null of no level shift, and consequently there is no adjustment. The SD approach differs more noticeably from the raw series, primarily due to adjustments for slow, downward drifts in the series during the mid-1990s and mid-2000s.

4. The predictive power of aggregate illiquidity

This section considers whether aggregate illiquidity measures contain information that is helpful for forecasting stock returns and various measures of economic activity.

4.1. Model specification and inference issues

Our initial tests are based on predictive regressions that incorporate lagged illiquidity measures. The predictive regressions that form the basis of our tests take the form

$$y_{t+1} = \alpha + \sum_{j=1}^{J} \theta_j y_{t+1-j} + \beta x_t + \gamma' z_t + \epsilon_{t+1}, \tag{9}$$

where y_{t+1} is the economic variable of interest (e.g., excess stock returns or a measure of economic activity), x_t indicates a particular aggregate illiquidity measure, and z_t represents a vector-valued set of additional macroeconomic and financial variables included in the forecasting model. The model potentially includes lags of the dependent variable to accommodate serial correlation. This feature is more relevant for measures of economic activity, which clearly display serial correlation. Under the null hypothesis of no Granger causality, $\beta = 0$. This hypothesis may be tested in a standard regression framework. For the monthly stock return regressions that serve as our initial focus, we set $J = 0$. Our initial tests focus on regressions that do not include additional macroeconomic variables such as bond spreads or stock return volatility ($z_t = 0$ in Eq. (9)).

The high degree of persistence associated with various illiquidity measures complicates statistical inference in the linear predictive regression model of Eq. (9). Persistent regressors can cause econometric problems in predictive regressions, particularly when there is correlation between innovations in the target variable and predictor (Stambaugh, 1999). OLS estimates of slope coefficients in such regressions can be biased and/or inconsistent. In addition, the t-statistic testing the null hypothesis of no Granger causality converges to a nonstandard asymptotic distribution when the regressor follows a local to unity process (Campbell and Yogo, 2006; Jansson and Moreira, 2006). The illiquidity measures we construct are better characterized as long memory processes as opposed to local to unity processes. However, long memory regressors generate similar econometric problems in predictive regressions (Maynard and Phillips, 2001; Maynard et al., 2013).

To assess the severity of inference problems associated with predictive regressions involving aggregate illiquidity measures, we conduct a Monte Carlo simulation analysis. The simulation analysis focuses on the finite sample bias of slope coefficient estimates and the size associated with Granger causality tests. We simulate data under the null of no predictability according to the following time-series model:

$$\begin{aligned} y_t - \mu_y &= \epsilon_{1,t}, \\ (1-L)^d (x_t - \mu_x) &= \epsilon_{2,t}, \\ \epsilon_t \equiv (\epsilon_{1,t}, \epsilon_{2,t})' &\sim \text{i.i.d. } N(0, \Sigma), \end{aligned} \tag{10}$$

where Σ is a symmetric, positive definite matrix with (potentially) nonzero off-diagonal elements. The contemporaneous correlation between shocks is $\rho = \Sigma_{12}/\sqrt{\Sigma_{11}\Sigma_{22}}$. The forecasting variable x potentially exhibits long memory behavior via the parameter d. The simulation analysis considers alternative ρ values ranging from -0.95 to 0.95, and values of the fractional differencing parameter d ranging from zero to one. Remaining parameters in the system (10) are calibrated based on monthly data for excess stock returns (y_t) and the *fht* illiquidity measure (x_t) over the sample period 1948–1997.[16] We choose a sample size (T) of 480 (40 years of monthly data) and report Monte Carlo results based on 2500 simulations.

Table 3 presents results. Panel A of Table 3 reports the simulation-based estimate of the bias associated with the OLS estimate of $\hat{\beta}$ in a predictive regression of the form $y_{t+1} = \alpha + \beta x_t + \nu_{t+1}$. Panel B reports the bias as a percentage of the simulation-based root mean square error (RMSE). This helps convey the economic significance of the bias associated with β. Finally, Panel C shows the simulated size of a two-sided t-test of the null that $\beta = 0$ based on the OLS estimate of β at the 5% level. Consistent with theory, the Monte Carlo results indicate essentially no bias when contemporaneous shocks to y_t and x_t are uncorrelated. As expected, the OLS estimate of β is negatively (positively) biased when $\rho > 0$ ($\rho < 0$). The magnitude of the bias increases as d increases from zero to 0.5. As d increases from 0.5 to 1.0 (the unit root case), the magnitude of the bias decreases. This is due to the fact that the simulation design holds fixed the variance of shocks ϵ_t across cases considered, but not the variance of x_t. In fact, for $d > 0.5$ the variance of x_t explodes similar to the unit root case. Upon normalizing by the RMSE (Panel B), the severity of the bias steadily increases as d rises from zero to unity. The bias is economically significant when there is long memory and when substantial correlation exists between contemporaneous shocks to y_t and x_t.

Granger causality tests are oversized for nonzero values of ρ, particularly in combination with larger values of d

[16] This sample period is chosen to avoid artificial inflation of persistence parameters due to structural shifts around tick-size reductions in 1997 and 2001. It is straightforward to generalize the system of Eq. (10) to permit additional short memory dynamics in x_t. In unreported additional analyses, we obtain qualitatively similar results upon including an AR(1) component in x_t that is calibrated to estimates from the data.

Table 3
Finite-sample bias in predictive regressions with long memory variables.

This table presents Monte Carlo evidence regarding the properties of the OLS estimator of the slope coefficient β in the predictive regression $y_{t+1} = \alpha + \beta x_t + \nu_{t+1}$. Data are generated under the null hypothesis of no predictability ($\beta = 0$) for the following system:

$$y_t - \mu_y = \epsilon_{1,t},$$
$$(1-L)^d(x_t - \mu_x) = \epsilon_{2,t},$$
$$\epsilon_t \equiv (\epsilon_{1,t}, \epsilon_{2,t})' \sim \text{i.i.d. } N(0, \Sigma),$$

where d represents the long memory parameter and Σ is a symmetric, positive definite matrix with (potentially) nonzero off-diagonal elements. The contemporaneous correlation between shocks is $\rho = \Sigma_{12}/\sqrt{\Sigma_{11}\Sigma_{22}}$. Simulation results are presented for specified values of d and ρ. Remaining parameters are calibrated based on monthly data for excess stock returns (y_t) and the *fht* illiquidity measure (x_t) over the sample period 1948–1997. Panel A presents the Monte Carlo estimate of the bias associated with $\hat{\beta}$, the OLS estimate of the slope coefficient in the predictive regression (multiplied by 100 for readability). Panel B shows the magnitude of the bias, expressed as a percentage of the root mean square error (RMSE) associated with the estimator, where the RMSE is computed using the Monte Carlo results. Panel C addresses the size of the standard (two-sided) t-test of the null hypothesis that $\beta = 0$ based on the OLS estimator. Specifically, Panel C reports the fraction of simulations in which the null hypothesis is rejected at the 5% level. Results are based on 2,500 simulations.

d/ρ	−0.95	−0.9	−0.8	−0.4	−0.2	0	0.2	0.4	0.8	0.9	0.95
Panel A: Bias											
0.0	0.00	0.06	0.05	0.05	−0.05	−0.12	−0.01	−0.01	−0.15	−0.12	−0.12
0.2	0.45	0.42	0.31	0.22	0.07	−0.01	0.05	−0.20	−0.44	−0.36	−0.51
0.4	0.88	0.85	0.70	0.41	0.21	0.03	−0.26	−0.44	−0.87	−0.87	−0.94
0.5	1.24	1.11	1.03	0.56	0.23	0.02	−0.22	−0.47	−1.06	−1.26	−1.24
0.6	1.36	1.26	1.13	0.60	0.28	−0.02	−0.25	−0.57	−1.16	−1.27	−1.37
0.8	1.14	1.10	0.95	0.46	0.26	−0.01	−0.23	−0.46	−0.94	−1.09	−1.18
1.0	0.63	0.61	0.54	0.27	0.14	−0.01	−0.15	−0.27	−0.50	−0.61	−0.65
Panel B: Absolute value of bias as % of RMSE											
0.0	0.17	2.20	1.91	1.69	1.71	4.17	0.46	0.38	5.52	4.61	4.49
0.2	16.77	15.18	11.48	8.22	2.48	0.29	1.83	7.24	16.07	13.22	18.54
0.4	37.27	35.89	29.40	17.28	9.43	1.15	11.42	19.06	36.19	36.50	38.85
0.5	54.37	50.39	46.89	27.61	11.68	0.85	11.20	23.62	48.51	54.05	54.47
0.6	66.23	62.76	58.20	35.75	17.25	1.26	15.39	34.20	58.79	62.90	65.52
0.8	76.51	74.75	70.53	45.83	27.95	1.68	24.65	45.90	70.49	74.20	76.59
1.0	76.58	74.64	72.75	52.35	32.17	1.76	32.89	53.09	71.40	75.46	75.31
Panel C: Empirical size of t-test (%)											
0.0	5.92	5.00	5.24	5.44	4.96	5.40	4.64	5.32	5.44	4.36	5.28
0.2	5.20	5.48	5.20	4.92	5.28	5.48	5.36	5.24	5.92	5.88	5.32
0.4	5.40	5.84	6.04	6.00	4.68	5.36	5.40	5.28	6.32	5.84	5.96
0.5	7.64	6.36	7.04	6.28	5.24	5.08	5.32	5.84	6.72	8.24	7.68
0.6	9.60	9.16	8.44	6.40	6.32	5.12	6.04	6.04	8.40	9.16	10.64
0.8	18.08	17.04	14.00	7.56	6.64	5.32	6.12	7.56	13.92	17.16	18.40
1.0	27.36	25.52	19.76	9.24	5.80	5.04	6.32	9.20	20.20	25.92	28.24

(Panel C). The case $d = 1$ and $\rho = -0.95$ produces rejections of the null approximately 27% of the time. This is consistent with Stambaugh (1999) and related papers focusing on the dividend yield and similar financial ratios. Although the frequency of rejections falls toward 5% as the value of d decreases from one to zero, the t-tests continue to be noticeably oversized whenever $d > 0.4$ and $|\rho| > 0.2$.

The simulation results in Table 3 indicate that biased coefficients and oversized Granger causality tests are potential concerns, particularly for stock return forecasting regressions.[17] As an alternative to conventional inference methods, we employ a bootstrap procedure analogous to that frequently used in exchange rate and stock return forecasting regressions involving near unit-root predictors (Mark, 1995; Killian, 1999). The bootstrap is implemented as follows for stock return forecasting regressions. First, we estimate the predictive system under the null hypothesis of no predictability, using the full sample of observations. This yields a time series of estimated shocks $\hat{\epsilon}_t$. To capture potential short memory dynamics in illiquidity measures, we generalize the system of Eq. (10) to incorporate an AR(1) component for x_t under the null. We then generate bootstrapped time series under the null by drawing with replacement from the estimated residuals $\hat{\epsilon}_t$. This procedure preserves the dependence structure of the predictor variable, including potential long memory behavior, as well as the contemporaneous correlation between shocks $\epsilon_{1,t}$ and $\epsilon_{2,t}$. The Internet Appendix provides details and an extension to persistent targets such as output growth. We report bootstrap-based p-values for one- and two-sided tests of the null that $\beta = 0$ in Eq. (9). In addition, we report a bias-corrected estimate of β, defined as $\hat{\beta}^* \equiv \hat{\beta} - (1/B)\sum_{i=1}^{B}\hat{\beta}_i^*$, where $\hat{\beta}$ is the OLS estimate of β and $\hat{\beta}_i^*$ indicates bootstrap estimates of β under the null of no predictability.

[17] Empirical estimates of ρ reported in Table 1 for various illiquidity measures tend to cluster around a value of approximately −0.3 when the target variable is stock returns, but are close to zero when the target variable is output growth.

4.2. In-sample predictive regressions for stock returns

Table 4 presents results for one-month-ahead univariate stock return forecasting regressions. The table shows estimated slope coefficients, robust standard errors with associated one- and two-sided p-values, and the regression R^2-value. For the one-sided test, the null hypothesis is $\beta \leq 0$. The right-hand portion of the table shows bias-corrected slope estimates and bootstrap-based p-values based on the bootstrap procedure discussed in Section 4.1. Different panels display results for alternative adjustments to the underlying illiquidity measures. Panel A presents results for unadjusted measures, and subsequent panels display results for measures that are volatility- and/or break-adjusted using various approaches to break adjustment.

Most raw illiquidity measures fail to successfully forecast excess market returns (Panel A). Break adjustment strengthens the case for stock return predictability, as several break-adjusted illiquidity measures are statistically significant predictors of returns. The best-performing predictor is the break-adjusted zeros measure, which is highly significant under both standard and bootstrap-based inference procedures. There is also some evidence that other break-adjusted measures including the *fht*, *tick*, and *ami* measures forecast returns. Results are generally similar under the FS and RT break-adjustment methods (Panels B and C). The evidence for predictability is weaker under the SD approach (Panel D), but the stochastically detrended *zeros* measure remains significant.

The final two panels illustrate the effects of decomposing illiquidity measures into a volatility component (*vol*) and a residual, in addition to adjusting for breaks using either the RT (Panel E) or SD (Panel F) approach. Volatility-adjustment further strengthens the evidence for stock return forecasting ability for most illiquidity measures.[18] The break- and volatility-adjusted measures *fht*, *tick*, *zeros*, and *ami* are all highly significant, irrespective of whether inference is based on conventional or bootstrap methods, under the RT break-adjustment approach. In addition, the R^2-values associated with these variables are in excess of 1.1% and compare favorably to in-sample R^2-values for a large body of popular stock return forecasting variables considered by Goyal and Welch (2008). The slope coefficients associated with all of these variables are positive, so that an increase in volatility-adjusted illiquidity is associated with higher expected returns, in accordance with theory. As with earlier results, break adjustment via the SD approach yields weaker results; however, the volatility-adjusted *fht* and *tick* measures remain significant (along with *zeros*). The Internet Appendix presents results for volatility- and break-adjusted measures using the FS break-adjustment approach. These are similar to results under the RT approach, and in some cases additional illiquidity measures such as *cs* and *amito* become marginally significant. The Internet Appendix also shows that we obtain similar results using alternative stock market return proxies.

The improved stock return-forecasting performance of volatility-adjusted measures is driven by the fact that the estimated return-forecasting coefficient associated with aggregate volatility (*vol*) is negative. The slope coefficients for unadjusted illiquidity measures are essentially weighted averages of return-forecasting coefficients for aggregate volatility and the corresponding volatility-adjusted illiquidity measure. The negative return-forecasting coefficient for volatility pulls the return-forecasting coefficients for unadjusted illiquidity measures closer to zero, and removing this component reveals a stronger predictive relation.

Bias-corrected bootstrap slope estimates are typically smaller than OLS-based estimates for raw illiquidity measures (Panel A), indicating some manifestation of the "Stambaugh bias" in our setting. However, the severity of the bias declines markedly for volatility- and break-adjusted measures (Panels E and F), and inference results tend to be similar under the OLS approach relative to the bootstrap method. A feature that mitigates the bias in our setting relative to predictive regressions involving financial ratios is the relatively modest correlation between contemporaneous shocks to stock returns and illiquidity measures. In addition, volatility-adjusted measures behave similarly to the *zeros* measure, and shocks to this measure are nearly uncorrelated with stock returns (see Table 1). This helps explain why bias-corrections to slope estimates are more pronounced for raw and break-adjusted measures relative to volatility-adjusted measures.

Fig. 4 illustrates the consistency of evidence of stock return-forecasting ability for key break- and volatility-adjusted illiquidity measures over alternative sample periods. The top panel of Fig. 4 plots t-statistics associated with OLS slope coefficient estimates for predictive regressions for excess stock returns using a 30-year rolling window. The initial estimation sample is 1927.2–1957.1. The bottom panel of Fig. 4 plots the R^2-values (as percentages) associated with these rolling predictive regressions. We present results for four volatility- and break-adjusted measures (using the RT approach) for which we obtain strong evidence of stock return predictability over the base sample period 1948–2015. The rolling window results in Fig. 4 incorporate additional data prior to 1948 along with illustrating performance over a rolling sub-period.

Rolling t-statistics are positive (as expected) across all sub-periods and are generally significant at the 5% level from the 1970s onward, with the exception of a period from the turn of the century until the onset of the financial crisis. Likewise, the corresponding R^2-values are generally in excess of 1% over this period. The weakest evidence for predictability occurs over the earliest sub-periods we examine; however, several predictors are significant for sub-periods ending in the mid-1960s. We offer two final remarks regarding the rolling regression results. First, many

[18] An alternative to conducting a predictive regression using a volatility-adjusted illiquidity measure is to include volatility as an additional forecasting variable in a bivariate predictive regression. The Frisch–Waugh–Lovell theorem implies that the OLS slope coefficient on the illiquidity measure in the bivariate regression is identical to that obtained by regressing the residuals from a projection of the forecasting target on volatility on the residuals from a projection of the illiquidity measure on volatility. Our estimates differ slightly from inference based on a bivariate regression because 1) we adjust most illiquidity measures for volatility using the analytical approach rather than an OLS projection; and 2) we do not first project the forecasting target onto volatility. We obtain similar results from bivariate regressions; however, and subsequent analyses explicitly include volatility in multivariate predictive regressions.

Table 4
Forecasting regressions for stock returns: monthly frequency.

The table presents results for one-month-ahead predictive regressions for excess returns on the NYSE index using various measures of aggregate illiquidity. The table reports results from regressions of the form

$$ret_{t+1} = \alpha + \beta x_t + \epsilon_{t+1},$$

where ret_{t+1} is the monthly NYSE excess return and x_t represents the specified forecasting variable. For each forecasting variable, the table presents OLS and bootstrap estimates. The presented OLS estimates include the slope coefficient $\hat{\beta}$, the associated Newey–West standard error, a one-sided p-value (p-val$_{1s}$), a two-sided p-value (p-val$_{2s}$), and an R^2-value. The presented bootstrap estimates include the slope coefficient $\hat{\beta}$ and the associated one-sided p-value (p-val$_{1s}$) and two-sided p-value (p-val$_{2s}$). The sample period is 1948–2015. See Section 2 of the paper or the Internet Appendix for variable definitions.

	OLS					Bootstrap		
	$\hat{\beta}$	SE	p-val$_{1s}$	p-val$_{2s}$	R^2	$\hat{\beta}^*$	p-val$_{1s}$	p-val$_{2s}$
Panel A: Raw illiquidity measures								
roll	−0.22	0.85	0.60	0.80	0.02	−0.31	0.75	0.64
cs	0.26	0.85	0.38	0.76	0.03	0.09	0.35	0.44
fht	0.17	0.16	0.14	0.29	0.14	0.09	0.29	0.39
tick	0.19	0.17	0.12	0.24	0.17	0.08	0.26	0.28
zeros	0.17	0.18	0.17	0.33	0.16	0.22	0.16	0.42
ami	0.38	0.16	0.01	0.02	0.59	0.30	0.03	0.03
amito	0.16	0.30	0.30	0.60	0.04	0.01	0.45	0.57
PS	−6.55	6.58	0.84	0.32	0.20	−6.80	0.91	0.20
hm	−0.48	0.82	0.72	0.56	0.09	−0.36	0.70	0.49
Panel B: Break-adjusted illiquidity measures (FS approach)								
roll	−0.15	0.86	0.57	0.86	0.01	−0.24	0.68	0.76
cs	0.11	0.72	0.44	0.88	0.01	−0.04	0.50	0.71
fht	0.91	0.66	0.09	0.17	0.41	0.77	0.03	0.03
tick	0.98	0.61	0.05	0.11	0.50	0.78	0.02	0.02
zeros	2.59	0.91	0.00	0.00	1.13	2.70	0.00	0.00
ami	0.57	0.24	0.01	0.02	0.84	0.49	0.00	0.00
amito	0.32	0.64	0.31	0.62	0.05	0.14	0.31	0.36
PS	−5.37	6.84	0.78	0.43	0.13	−5.71	0.87	0.29
hm	−1.10	1.00	0.86	0.27	0.25	−1.02	0.90	0.17
Panel C: Break-adjusted illiquidity measures (RT approach)								
roll	−0.23	0.84	0.61	0.78	0.03	−0.34	0.75	0.64
cs	0.21	0.74	0.39	0.77	0.02	0.04	0.43	0.59
fht	0.82	0.59	0.08	0.16	0.34	0.68	0.04	0.04
tick	0.87	0.48	0.04	0.07	0.43	0.69	0.02	0.02
zeros	2.28	0.91	0.01	0.01	1.16	2.37	0.00	0.00
ami	0.49	0.22	0.01	0.02	0.67	0.41	0.01	0.01
amito	0.18	0.49	0.36	0.71	0.02	0.00	0.48	0.59
PS	−6.00	6.40	0.83	0.35	0.17	−6.12	0.89	0.23
hm	−0.53	1.04	0.69	0.61	0.07	−0.43	0.73	0.47
Panel D: Break-adjusted illiquidity measures (SD approach)								
roll	−0.96	0.95	0.84	0.32	0.32	−1.02	0.95	0.11
cs	−1.20	1.25	0.83	0.34	0.27	−1.40	0.99	0.06
fht	0.69	1.04	0.25	0.51	0.14	0.56	0.10	0.13
tick	0.48	1.04	0.32	0.65	0.06	0.26	0.22	0.23
zeros	1.78	0.88	0.02	0.04	0.59	1.86	0.00	0.01
ami	−0.09	0.45	0.58	0.85	0.01	−0.20	0.83	0.68
amito	−0.26	0.69	0.65	0.71	0.03	−0.43	0.86	0.53
PS	−7.65	6.63	0.88	0.25	0.25	−7.81	0.92	0.16
hm	0.31	1.07	0.39	0.77	0.02	0.39	0.33	0.72
Panel E: Decomp. of illiquidity into volatility and residual components (RT approach)								
vol	−0.67	0.86	0.78	0.44	0.17	−0.83	0.97	0.15
roll	0.33	0.87	0.35	0.71	0.03	0.29	0.36	0.68
cs	0.68	0.67	0.16	0.31	0.15	0.77	0.16	0.38
fht	2.18	0.68	0.00	0.00	1.45	2.28	0.00	0.00
tick	1.82	0.63	0.00	0.00	1.36	1.84	0.00	0.00
zeros	2.28	0.91	0.01	0.01	1.16	2.38	0.00	0.01
ami	0.70	0.22	0.00	0.00	1.13	0.64	0.00	0.00
amito	0.54	0.58	0.18	0.35	0.21	0.49	0.10	0.16
PS	−5.78	6.51	0.81	0.37	0.14	−6.06	0.86	0.28
hm	−0.97	0.98	0.84	0.32	0.18	−0.85	0.83	0.27
Panel F: Decomp. of illiquidity into volatility and residual components (SD approach)								
vol	−0.67	0.86	0.78	0.44	0.17	−0.82	0.97	0.16
roll	−0.83	1.03	0.79	0.42	0.11	−0.85	0.83	0.35
cs	−0.46	1.34	0.63	0.73	0.02	−0.40	0.65	0.68
fht	1.83	0.85	0.02	0.03	0.72	1.89	0.00	0.01
tick	1.70	0.96	0.04	0.08	0.54	1.71	0.00	0.01
zeros	1.78	0.88	0.02	0.04	0.59	1.89	0.00	0.01
ami	0.12	0.45	0.40	0.79	0.01	0.06	0.40	0.63
amito	0.32	0.69	0.32	0.64	0.03	0.25	0.32	0.55
PS	−7.33	6.68	0.86	0.27	0.21	−7.28	0.90	0.20
hm	0.02	1.09	0.49	0.99	0.00	0.07	0.48	0.98

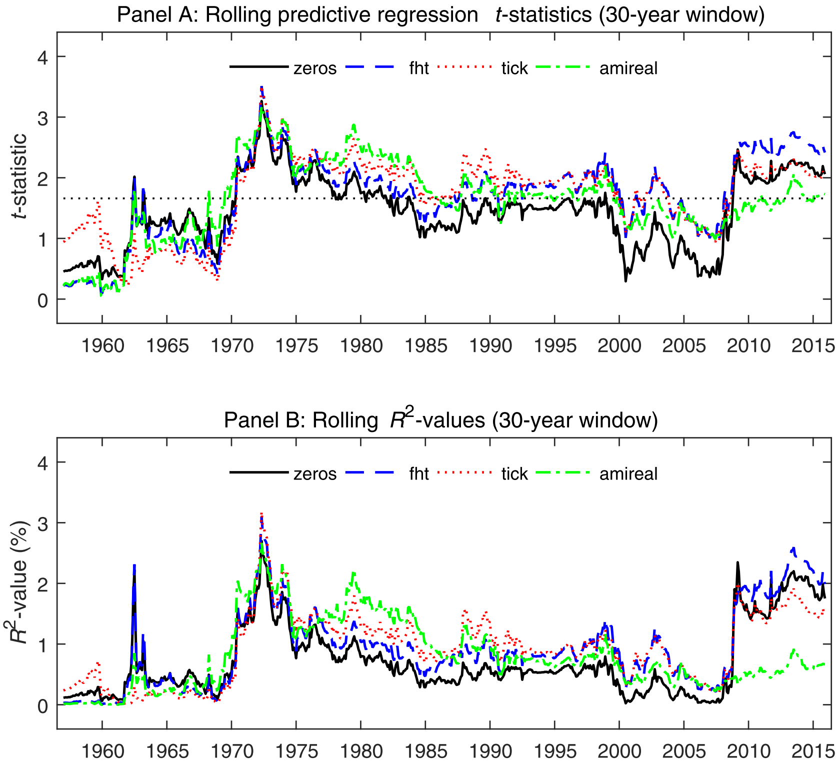

Fig. 4. This figure displays t-statistics and R^2-values for rolling predictive regressions of excess stock market returns on lagged illiquidity measures. Results are reported for four break- and volatility-adjusted illiquidity measures: *zeros*, *fht*, *tick*, and *ami*. Break-adjustment is based on the real time (RT) approach. Panel A plots t-statistics associated with OLS estimates of the slope coefficient $\hat{\beta}$ in univariate predictive regressions of excess NYSE returns (*ret*) for selected break- and volatility-adjusted illiquidity measures. The reported t-statistics are based on Newey–West standard errors. A thin, dotted horizontal line indicates the cutoff of 1.66 associated with statistical significance at the 10% level for a standard two-sided hypothesis test of the null of no predictability. Panel B plots rolling R^2-values (as percentages) for the same predictive regressions. The rolling regressions employ an estimation window of 30 years, with the initial estimation sample beginning in 1927.2 and ending in 1957.1.

traditional stock return-forecasting variables appear to break down in recent decades (see, e.g., Goyal and Welch, 2008). Although results in Fig. 4 indicate the onset of a similar breakdown for the adjusted illiquidity measures in the early 2000s, both rolling t-statistics and R^2-values are near historical *maximums* over the most recent 30-year period, due largely to strong performance during the financial crisis and its aftermath. Second, t-statistics and R^2-values tend to increase during several other periods associated with economic recessions, including the mid-1970s and (to a lesser extent) during the early 1980s and 1990s.

It is possible that apparent forecasting power associated with illiquidity measures simply reflects the fact that these measures are correlated with other financial variables frequently employed in stock return-forecasting models. Table 5 presents results for predictive regressions for monthly excess stock market returns that include additional financial forecasting variables z_t. We consider three increasing sets of variables for z_t. The first, denoted $z_{t,1}$, includes the lagged excess return to capture any persistence in monthly returns as well as (lagged) *vol*. The second ($z_{t,2}$) adds several common bond-related forecasting instruments, including the (stochastically detrended) short-term interest rate, the term spread, the default spread, and the commercial paper-to-Treasury spread. The final set ($z_{t,3}$) adds the book-to-market ratio, earnings-price ratio, and net equity issuance. Given the relatively high dimensionality of the predictors for these models, we report OLS estimates of $\hat{\beta}$, associated t-statistics based on Newey–West standard errors, and ΔR^2-values, defined as the increase in R^2 associated with including the illiquidity measure x_t relative to an otherwise identical specification that excludes it.

Results in Table 5 illustrate that many break- and volatility-adjusted illiquidity measures remain statistically significant predictors of stock returns upon including additional return-forecasting variables. For adjusted illiquidity measures based on the RT method, slope coefficients and ΔR^2-values associated with predictors remain relatively stable across the alternative specifications. In some cases, the inclusion of additional forecasting instruments actually strengthens the case for predictability for several illiquidity measures, particularly for those based on the stochastic detrending (SD) procedure (Panel B of Table 5). Interestingly, the pattern of evidence is relatively similar under the RT and SD adjustment approaches after including the full set of controls ($z_{t,3}$).

Table 5
Multivariate return-forecasting regressions.
The table presents one-month-ahead predictive regressions for excess returns on the NYSE index (*ret*) in which other common return-forecasting instruments are included in the forecasting model, along with the illiquidity measures of central interest in this paper. The predictive regression model takes the form:

$$ret_{t+1} = \alpha + \beta x_t + \gamma' z_t + \epsilon_{t+1},$$

where ret_{t+1} denotes excess stock returns and x_t indicates a break- and volatility-adjusted aggregate illiquidity measure of interest. The vector z_t represents an additional set of macroeconomic and financial forecasting variables included in the model. We consider three choices for z_t. z_1 consists of lagged volatility (*vol*) and two lags of the dependent variable. z_2 consists of z_1 as well as the lagged term spread, default spread, commercial paper-to-Treasury spread, and stochastically detrended T-bill rate. Finally, z_3 consists of z_2 as well as the lagged book-to-market ratio, earnings-price ratio, and net equity issuance. For each forecasting variable, the table presents the estimated slope coefficient $\hat{\beta}$, the associated t-statistic based on Newey–West standard errors, and ΔR^2, defined as the increase in R^2-value relative to a similar regression that excludes x_t (the above model with $\beta = 0$). Panel A presents results for break- and volatility-adjusted illiquidity measures using the real time (RT) approach. Panel B presents results for break- and volatility-adjusted measures using the stochastic detrending (SD) approach. The sample period is 1948–2015. See Section 2 of the paper or the Internet Appendix for variable definitions.

	Controls $= z_1$			Controls $= z_2$			Controls $= z_3$		
	$\hat{\beta}$	t-stat	ΔR^2	$\hat{\beta}$	t-stat	ΔR^2	$\hat{\beta}$	t-stat	ΔR^2
Panel A: Break- and volatility-adjusted illiquidity measures (RT approach)									
roll	0.29	0.33	0.02	0.08	0.08	0.00	0.04	0.05	0.00
cs	0.61	0.94	0.12	0.23	0.32	0.01	0.24	0.29	0.01
fht	2.01	3.15	1.16	2.06	3.02	1.15	1.74	2.34	0.77
tick	1.68	3.05	1.10	1.72	2.87	1.10	1.42	2.16	0.67
zeros	2.16	2.63	0.87	2.41	2.87	1.06	2.01	2.18	0.63
ami	0.69	3.23	1.08	0.69	3.04	1.04	0.55	2.12	0.60
amito	0.41	0.77	0.10	0.79	1.45	0.32	0.99	1.81	0.39
PS	−3.73	−0.60	0.06	−0.73	−0.12	0.00	−1.25	−0.19	0.01
hm	−1.70	−1.87	0.47	−1.50	−1.69	0.35	−1.04	−1.05	0.16
Panel B: Break- and volatility-adjusted illiquidity measures (SD approach)									
roll	−0.80	−0.77	0.11	−0.76	−0.74	0.09	−0.78	−0.77	0.10
cs	−0.97	−0.74	0.07	−1.22	−0.89	0.10	−0.84	−0.59	0.04
fht	1.55	1.69	0.48	2.19	2.32	0.78	2.43	2.60	0.90
tick	1.43	1.51	0.36	1.94	1.95	0.54	2.12	2.10	0.62
zeros	1.49	1.60	0.37	2.35	2.40	0.74	2.46	2.56	0.78
ami	0.17	0.39	0.02	0.48	1.11	0.15	0.56	1.24	0.19
amito	0.20	0.30	0.01	0.69	0.99	0.11	0.97	1.40	0.22
PS	−4.50	−0.71	0.07	−1.81	−0.28	0.01	−1.54	−0.24	0.01
hm	−1.13	−1.01	0.12	−1.03	−0.93	0.10	−0.73	−0.64	0.05

4.3. Out-of-sample forecasting performance

Many recent papers in the stock return predictability literature assess predictive ability using an out-of-sample research design. There remains contention in the literature regarding whether in- or out-of-sample model comparison is preferred. Without taking a stand on this issue, this section provides an out-of-sample analysis to complement earlier in-sample results.[19]

The out-of-sample analysis compares the performance of excess return forecasts based on predictive regressions involving illiquidity measures to benchmark forecasts. More specifically, let y_{t+1} represent the forecasting target (excess stock returns in the present analysis) and $\hat{y}_{t+1}$ denote a one-step-ahead forecast of y_{t+1} from a predictive regression using only data available at time t in order to simulate a 'real time' forecasting environment. The corresponding sample mean square prediction error (MSPE) for the predictive regression is

$$\hat{\sigma}^2 \equiv P^{-1} \sum (y_{t+1} - \hat{y}_{t+1})^2, \tag{11}$$

where P denotes the number of out-of-sample forecasts. Similarly, let $\hat{y}^{BMK}_{t+1}$ denote the one-step-ahead forecast based on the benchmark model, with corresponding sample MSPE equal to $\hat{\sigma}^2_{BMK}$. We report the out-of-sample R^2 statistic (Campbell and Thompson, 2008; Goyal and Welch, 2008). This statistic is defined as

$$R^2_{OOS} = 1 - \frac{\hat{\sigma}^2}{\hat{\sigma}^2_{BMK}}, \tag{12}$$

expressed as a percentage. The R^2_{OOS} statistic serves as the out-of-sample analog of the ΔR^2 statistic reported for in-sample regressions. Following most literature concerning stock return predictability, the benchmark forecast for stock market returns is the historical average. Forecasting models are estimated using either a rolling or recursive (expanding) procedure with an initial estimation sample of 30 years of data. We consider two alternative out-of-sample evaluation periods: 1957–2015 and 1978–2015.

Although it is informative, the R^2_{OOS}-value represents a particular sample realization. Hence, it is of interest to test

[19] Out-of-sample designs seek to address potential concerns regarding structural instability and data mining. Ferson et al. (2003) show that data mining can exacerbate spurious regression bias in predictive regressions involving persistent predictors. For current perspectives regarding in-sample versus out-of-sample inference, see Diebold (2015) and published commentary on this article in the same issue.

whether differences in forecast performance are statistically significant. Statistical inference associated with out-of-sample forecast performance is nuanced when forecasts rely on estimated parameters (West, 1996). We conduct two alternative tests for differences in out-of-sample forecast performance. The first tests the null hypothesis of no Granger causality ($\beta = 0$ in Eq. (9)) in an out-of-sample design. A variety of approaches to inference exist in the literature [see Clark and McCracken (2011) for a comprehensive review]. We opt to apply a conceptually simple approach proposed by Clark and West (2007). The test is based on the quantity

$$CW = \hat{\sigma}^2_{BMK} - \hat{\sigma}^2 + \underbrace{P^{-1}\sum(\hat{y}_{t+1} - \hat{y}^{BMK}_{t+1})^2}_{\text{Adjustment}}. \quad (13)$$

The final term in Eq. (13) captures an adjustment for additional noise associated with the larger model's forecast under the null hypothesis. The test statistic associated with CW is not asymptotically normal; however, simulation evidence in Clark and West (2007) indicates that standard normal critical values lead to tests with actual sizes close to nominal size. The test is one-sided, because, under the alternative, $\sigma^2 < \sigma^2_{BMK}$.

We also apply the test for superior predictive ability proposed by Giacomini and White (2006). The Giacomini and White (2006) test focuses on differences in finite sample predictive ability, i.e., predictive ability incorporating the impact of parameter estimation on forecast performance. The distinction is important, because a variable may Granger cause stock returns, but fail to forecast returns more accurately than the benchmark forecast when parameters are estimated using finite samples, due to the bias-variance tradeoff. The Giacomini and White (2006) test statistic for equal (unconditional) predictive ability is:

$$\text{Giacomini–White} = \frac{\hat{\sigma}^2_{BMK} - \hat{\sigma}^2}{\hat{\sigma}_P/\sqrt{P}}, \quad (14)$$

where $\hat{\sigma}_P$ is a heteroskedasticity and autocorrelation consistent (HAC) estimator of the asymptotic variance $\sigma^2_P = \text{var}\left[\sqrt{P}(\hat{\sigma}^2_{BMK} - \hat{\sigma}^2)\right]$. In contrast with the Clark-West test, the Giacomini–White test is two-sided.

In addition to 'univariate' forecasting approaches that include only a single predictor, we also consider a forecast combination approach. Rapach et al. (2010) show that forecast combination methods improve the out-of-sample performance of stock return forecasts based on multiple predictors. Following Rapach et al. (2010), we employ a simple combined forecast defined as the mean of the forecasts produced by various illiquidity measures considered.[20] This combined forecast is denoted *Comb.* in tabulations of results.

Table 6 presents results for one-month-ahead out-of-sample forecasts of excess stock returns. For each variable, the table shows the ΔR^2_{OOS} summarizing out-of-sample performance, as well as the *t*-statistic associated with the Clark and West (2007) and Giacomini and White (2006) tests. To conserve space, we focus on results for break- and volatility-adjusted illiquidity measures using the RT and SD approaches to mitigate concerns regarding look-ahead bias. For comparison, we also report out-of-sample forecasting results for *vol* and a variety of common alternative stock return-forecasting variables. We compute a combined forecast over the set of alternative forecasting variables for comparison with that based on alternative illiquidity measures.

Consistent with in-sample results, the out-of-sample (OOS) evidence indicates that a number of volatility- and break-adjusted illiquidity measures forecast excess stock market returns. Measures that performed best in-sample, including *fht*, *tick*, *zeros*, and *ami*, generally outperform the historical average benchmark out-of-sample. In addition, the Clark-West out-of-sample test for Granger causality typically rejects the null of no predictability for these variables. The simple forecast combination approach that pools information from all illiquidity measures also performs well. Similar to in-sample results, explicit break adjustment via the RT approach yields better performance relative to the stochastic detrending (SD) approach.

The relatively strong out-of-sample forecasting performance of the adjusted illiquidity measures is perhaps best appreciated in comparison to the performance of many standard return forecasting instruments in Panel C of Table 6. Consistent with the spirit of results in Goyal and Welch (2008), aggregate volatility and the other financial predictors generally fail to outperform the historical average. The performance differences are particularly strident over the more recent OOS evaluation period (1978–2015). Over this period, virtually all traditional forecasting variables underperform the benchmark, with R^2_{OOS}-values below -1% not uncommon. In contrast, the more successful illiquidity measures often generate R^2_{OOS}-values over 1%. Forecast combination noticeably improves the performance of predictions based on traditional variables, as in Rapach et al. (2010). However, it is quite interesting to compare the performance of the combined forecasts in Panels A and C. The former includes only various illiquidity measures, whereas the latter includes a host of popular return-forecasting variables, such as the earnings yield, term spread, and default spread. And yet the combined forecast based on illiquidity measures alone performs as well or better than that constructed from traditional measures across the various scenarios we consider.

Fig. 5 plots the cumulative out-of-sample MSE difference between one-month-ahead stock return forecasts based on predictive regressions for selected break- and volatility-adjusted illiquidity measures, and the historical average benchmark. This form of plot, popularized by Goyal and Welch (2008), illustrates the time-series behavior of forecast performance relative to the benchmark. An upward slope indicates periods in which forecasts based on the corresponding illiquidity measure outperform the benchmark, whereas a downward slope indicates periods in which the benchmark is more accurate. The out-of-sample analysis period is 1957–2015 with model estimates

[20] Rapach et al. (2010) also develop more sophisticated combination approaches that adapt over time based on forecast performance; however, the simple averaging approach performs relatively well in their application and for simplicity we focus on this approach.

Table 6
Out-of-sample stock return-forecasting performance.
This table presents out-of-sample forecasts for monthly data. *CW* is the *t*-statistic for the Clark-West test. ΔR^2_{OOS} is the out-of-sample R^2-value relative to the benchmark. *GW* is the *t*-statistic for the Giacomini and White (2006) test for equal predictive ability. The forecasting target is the excess stock market return on the NYSE. The benchmark forecasting model is the historical average. Forecasts are based on estimates using either a rolling window of 30 years or a recursive estimation procedure with an initial estimation sample of 30 years. The *GW* statistic is omitted for the recursive scheme since the asymptotic results do not support recursive estimation. Results are reported for out-of-sample periods of 1957–2015 and 1978–2015. Panel A presents results for break- and volatility-adjusted illiquidity measures using the real time (RT) approach. Panel B presents results for break- and volatility-adjusted illiquidity measures using the stochastic detrending (SD) approach. Panel C shows results for volatility and various other standard return-forecasting variables for comparison. *Comb.* denotes the combined forecast based on variables listed in the corresponding panel.

	Rolling window						Recursive estimation			
	1957–2015			1978–2015			1957–2015		1978–2015	
	ΔR^2_{OOS}	*CW*	*GW*	ΔR^2_{OOS}	*CW*	*GW*	ΔR^2_{OOS}	*CW*	ΔR^2_{OOS}	*CW*
Panel A: Break- and volatility-adjusted measures (RT)										
roll	−0.04	1.32	−0.07	−0.27	0.45	−0.38	−1.45	−1.09	−0.77	−0.60
cs	0.06	1.32	0.08	−0.34	0.77	−0.34	−0.45	−1.37	−0.67	−0.26
fht	1.31	2.58	1.46	1.64	2.44	1.46	0.96	2.50	1.59	2.46
tick	0.94	2.35	1.11	1.23	2.12	1.06	1.02	2.38	1.20	1.99
zeros	0.97	1.97	1.05	1.13	1.64	0.98	0.82	1.92	1.18	1.72
ami	1.02	2.83	1.47	0.84	2.22	1.13	0.62	2.26	0.65	1.94
amito	−0.14	0.05	−0.32	0.13	0.54	0.21	−0.03	0.16	−0.18	−0.24
PS	−0.98	0.44	−0.82	−0.57	−0.66	−1.30	−0.50	0.46	−0.18	0.14
hm	−0.42	−0.24	−0.79	−0.95	−0.98	−1.34	−0.27	−0.27	−1.27	−0.70
Comb.	0.87	2.17	1.64	0.72	1.82	1.29	0.48	2.03	0.63	2.03
Panel B: Break- and volatility-adjusted illiquidity measures (SD)										
roll	−0.29	0.46	−0.54	−0.49	−0.39	−0.86	−1.35	0.34	−0.24	−0.46
cs	−0.02	0.54	−0.06	−0.21	0.04	−0.49	−0.50	−0.66	−0.38	−1.24
fht	0.44	1.57	0.69	0.81	1.73	0.93	−1.01	−1.41	0.69	1.63
tick	0.40	1.57	0.65	0.78	1.73	0.94	−1.43	−1.68	0.44	1.35
zeros	0.41	1.59	0.76	0.70	1.68	0.99	−1.38	−1.21	0.35	1.13
ami	−0.30	0.10	−0.63	−0.41	−1.05	−1.42	−1.15	−0.55	−0.37	−1.80
amito	0.31	1.58	0.57	−0.06	0.57	−0.11	−1.07	−0.27	−0.74	−0.89
PS	−0.73	0.68	−0.62	−0.81	−0.62	−1.32	−0.53	0.83	−0.36	0.12
hm	−0.75	−1.45	−1.96	−0.66	−1.18	−1.49	−0.43	−1.31	−0.44	−1.25
Comb.	0.38	1.44	0.97	0.21	1.00	0.61	−0.50	−0.55	0.02	0.34
Panel C: Volatility and other financial variables										
vol	−0.69	−0.38	−0.87	−0.57	−0.09	−0.57	−0.07	−0.35	−0.26	−0.19
TERM	0.07	2.13	0.06	−0.31	1.32	−0.21	0.12	1.24	−1.20	0.88
DEF	−0.93	−0.12	−1.28	−1.33	−0.61	−1.54	0.03	0.41	−0.78	−0.73
CPSP	−0.34	0.68	−0.15	−1.04	−0.64	−1.25	−0.04	0.83	−0.73	−0.02
TBL	−0.34	2.77	−0.19	−1.28	1.43	−0.66	−0.99	0.29	−2.65	0.93
NTIS	−0.34	1.36	−0.36	−0.70	1.03	−0.59	−0.91	0.40	−0.85	0.29
BMKT	−1.07	0.12	−1.44	−0.61	−0.72	−1.26	−1.81	0.49	−0.53	−0.09
ep	−0.79	−0.06	−0.96	−0.79	−0.29	−0.76	−1.51	0.56	−0.50	0.38
dp	0.04	1.41	0.06	−0.84	−0.04	−1.03	−0.21	1.17	−1.11	0.67
Comb.	0.84	2.36	1.63	0.23	1.01	0.50	0.41	1.85	0.28	1.21

produced using a rolling estimation window with 30 years of initial data.

When break adjustment is based on the RT approach (top panel), stock return forecasts based on illiquidity measures outperform the benchmark over the out-of-sample analysis period. The plot nearly always slopes upward or is roughly flat. In addition, there are several periods during which the illiquidity measures produce much more accurate forecasts relative to the benchmark. Most notable is the period from the onset of the financial crisis onward. Although this period yields the greatest gains relative to the benchmark, illiquidity measures perform well during several other periods, including around October 1987 and during the early 1970s. Although the out-of-sample performance gains relative to the historical average are more modest under the SD approach, the pattern of performance gains around the financial crisis and in the 1970s remains.

4.4. Predictive power for economic activity

This section considers the predictive content of illiquidity for measures of economic activity. Prior research by Næs et al. (2011) indicates that raw illiquidity measures forecast economic activity, and our primary interest involves exploring relative contributions from the volatility and residual components of these measures.

Table 7 presents one-month-ahead forecasting results for two macroeconomic targets for the period 1948–2015: the growth rate in industrial production Δip and the change in unemployment rate Δue. The table presents results analogous to those for stock return-forecasting regressions in Table 4 with two modifications. First, we omit *p*-values for one-sided tests to conserve space. Second, because both Δip and Δue exhibit serial correlation, we include four lags in the benchmark specification ($J = 4$ in

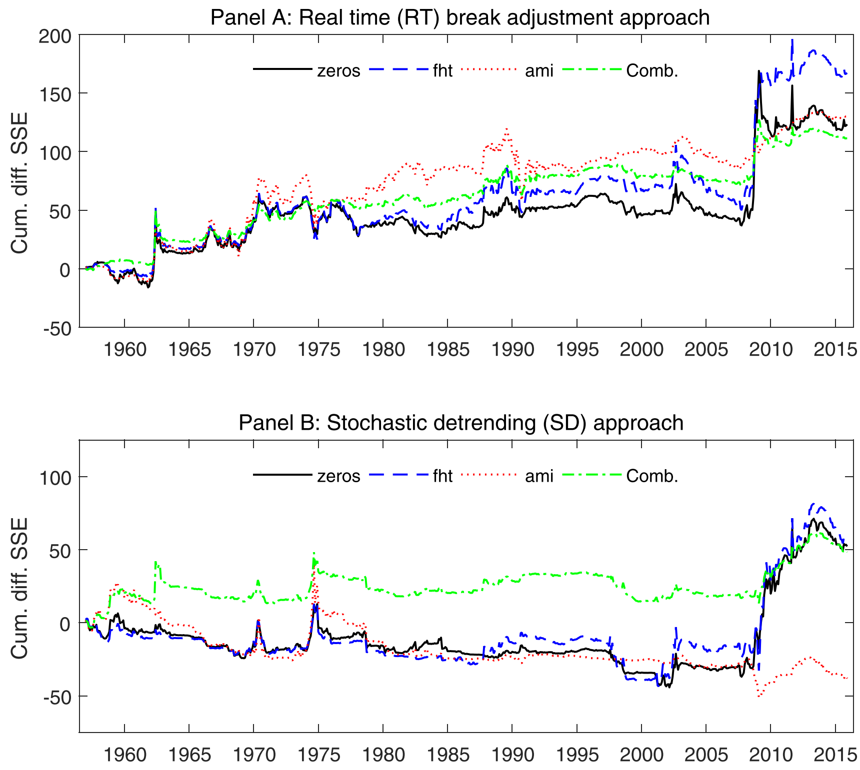

Fig. 5. This figure plots the out-of-sample cumulative difference in the sum of squared forecast error (Cum. diff. SSE), relative to the historical average benchmark, for one-month-ahead stock return forecasts for selected forecasts based on illiquidity measures. Forecasts for models incorporating illiquidity measures are based on Eq. (9), with no lags included ($J = 0$). The figure shows results for the following measures: *zeros*, *fht*, *ami*, and the simple forecast combination approach (*Comb.*) that averages forecasts in each period across all nine illiquidity measures. All illiquidity measures are volatility- and break-adjusted. Panel A presents results for the 'real time' (RT) break-adjustment method and Panel B shows results for the stochastic detrending (SD) approach. The initial out-of-sample forecast is based on parameters estimated using a rolling 30-year estimation sample beginning in 1927.2. Consequently, the initial out-of-sample forecast is for 1957:2 and the final out-of-sample forecast is for 2015.12.

Eq. (9)). Consequently, we report ΔR^2, defined as the increase in the R^2-value for the predictive regression relative to this benchmark univariate autoregressive model, expressed as a percentage.

The results in Table 7 indicate that aggregate illiquidity Granger causes economic activity. Among the break-adjusted illiquidity measures (Panel A), virtually all measures predict both economic activity target variables. Consistent with economic theory and intuition, higher illiquidity levels forecast lower future output growth (*ip*) and higher future unemployment (*ue*). The break-adjusted *hm* measure forecasts real activity; however, the direction of the relation differs from that of other illiquidity variables. This reflects the previously documented fact that aggregate price delay tends to decrease, rather than increase, in recessions. The incremental R^2 associated with the illiquidity measures is nontrivial, ranging from around 0.25% to 1.5% for specifications with significant predictors. Interestingly, the *zeros* measure does *not* appear to forecast economic activity, even after break-adjustment. Section 3.2 shows that the *zeros* measure can be interpreted as a spread measure adjusted for volatility. Consequently, the results in Panel A of Table 7 suggest that a significant portion of the forecasting power of illiquidity measures for real activity arises from the volatility component embedded in these measures.

Panel B of Table 7 provides results for the decomposition of aggregate illiquidity measures into a volatility component (*vol*) and a residual component. Aggregate volatility significantly forecasts growth in industrial production and growth in unemployment. The associated slope coefficients take the expected sign: high volatility predicts reduced economic activity, controlling for past activity. In contrast, slope coefficients associated with many volatility- and break-adjusted measures are statistically insignificant. There remains limited evidence of predictability associated with volatility-adjusted measures, however. For example, the volatility-adjusted *ami* measure forecasts growth in industrial production and unemployment. Bootstrap-based bias correction produces little change in slope coefficients relative to OLS estimates, and inference results are similar. The small magnitude of bias and similarity of inference results derive from the fact that shocks to macroeconomic growth measures are nearly uncorrelated with contemporaneous shocks to illiquidity measures. Table 8 shows that including additional forecasting variables strengthens the case for predictability, particularly for unemployment. Upon including the richest set of controls, six of the nine

Table 7
Forecasting regressions for macroeconomic activity: monthly frequency.
The table presents results for one-month-ahead predictive regressions for growth in industrial production (Δip) and growth in unemployment (Δue) using various measures of aggregate illiquidity. The table reports results from regressions of the form

$$y_{t+1} = \alpha + \sum_{j=1}^{J} \theta_j y_{t+1-j} + \beta x_t + \epsilon_{t+1},$$

where y_{t+1} is the dependent variable and x_t represents the specified forecasting variable. For each forecasting variable, the table presents OLS and bootstrap estimates. The presented OLS estimates include the slope coefficient $\hat{\beta}$, the associated Newey–West standard error, a two-sided p-value, and ΔR^2, defined as the increase in R^2-value relative to a similar regression that excludes x_t (the above model with $\beta = 0$). The presented bootstrap estimates include the slope coefficient $\hat{\beta}$ and the associated two-sided p-value. Four lags of the dependent variable are included in the predictive regression ($J = 4$). The sample period is 1948–2015. See Section 2 of the paper or the Internet Appendix for variable definitions.

	$y_{t+1} = \Delta ip$						$y_{t+1} = \Delta ue$					
	OLS				Bootstrap		OLS				Bootstrap	
	$\hat{\beta}$	SE	p-val	ΔR^2	$\hat{\beta}^*$	p-val	$\hat{\beta}$	SE	p-val	ΔR^2	$\hat{\beta}^*$	p-val
Panel A: Break-adjusted illiquidity measures (RT approach)												
roll	−0.42	0.12	0.00	1.46	−0.42	0.00	1.46	0.45	0.00	1.17	1.50	0.01
cs	−0.45	0.12	0.00	1.55	−0.45	0.00	1.57	0.43	0.00	1.22	1.58	0.00
fht	−0.27	0.11	0.02	0.62	−0.28	0.00	1.04	0.40	0.01	0.61	1.06	0.00
tick	−0.27	0.10	0.01	0.72	−0.27	0.00	1.07	0.35	0.00	0.73	1.09	0.00
zeros	0.22	0.18	0.21	0.21	0.20	0.18	0.00	0.73	1.00	0.00	−0.00	1.00
ami	−0.13	0.04	0.00	0.87	−0.14	0.00	0.58	0.16	0.00	1.03	0.60	0.00
amito	−0.26	0.08	0.00	0.87	−0.26	0.01	1.39	0.33	0.00	1.55	1.39	0.00
PS	−1.63	0.93	0.08	0.23	−1.59	0.16	7.42	3.80	0.05	0.30	7.32	0.10
hm	0.49	0.17	0.00	1.02	0.49	0.00	−1.61	0.67	0.02	0.72	−1.65	0.02
Panel B: Decomposition of illiquidity measures into volatility component and residual (RT approach)												
vol	−0.48	0.16	0.00	1.49	−0.48	0.00	1.32	0.66	0.05	0.75	1.32	0.01
roll	−0.17	0.14	0.23	0.14	−0.19	0.33	0.78	0.55	0.15	0.19	0.73	0.28
cs	−0.23	0.15	0.14	0.29	−0.25	0.20	0.65	0.55	0.24	0.16	0.67	0.38
fht	0.12	0.14	0.36	0.09	0.12	0.29	0.07	0.55	0.90	0.00	0.11	0.89
tick	0.04	0.12	0.72	0.01	0.02	0.75	0.25	0.50	0.61	0.03	0.34	0.66
zeros	0.22	0.18	0.21	0.21	0.20	0.21	0.00	0.73	1.00	0.00	0.09	1.00
ami	−0.09	0.05	0.07	0.30	−0.09	0.08	0.47	0.19	0.01	0.57	0.48	0.02
amito	−0.02	0.10	0.81	0.01	−0.03	0.80	0.63	0.37	0.09	0.34	0.63	0.11
PS	−1.17	0.96	0.22	0.11	−1.29	0.39	6.21	3.83	0.10	0.19	6.30	0.23
hm	0.38	0.16	0.02	0.49	0.37	0.07	−1.35	0.69	0.05	0.41	−1.30	0.10

adjusted illiquidity measures significantly forecast growth in unemployment. Associated ΔR^2-values for these measures range from roughly 0.3%–1.3%.

4.5. Principal components, quarterly data, and robustness checks

Our final results extend previous analyses in two directions. First, we construct a single measure from the set of underlying break- and volatility-adjusted illiquidity measures. To obtain this measure, we perform a standard principal component analysis using the full sample of (standardized) adjusted measures, and extract the resulting first principal component. The extraction of a global measure of illiquidity from our adjusted proxies is reminiscent of Korajczyk and Sadka (2008), who perform a similar exercise using standard (unadjusted) measures with application to pricing the cross-section of stock returns. Second, we examine predictability using quarterly, rather than monthly data. Many standard measures of economic activity, including GDP, are only available at the quarterly frequency. Analyzing quarterly data permits us to assess a variety of additional measures of activity, and provides a further robustness check on monthly results for excess stock returns. For expositional efficiency, we present results that combine these extensions, i.e., evidence from quarterly predictive regressions for both underlying illiquidity measures and the extracted principal component. The Internet Appendix provides additional monthly forecasting results based on a principal component extracted from monthly data.

Target variables at the quarterly frequency include excess stock market returns as well as the following measures of economic activity: 1) the growth rate in GDP Δgdp, 2) the growth rate of investment Δinv, 3) the growth rate of consumption $\Delta cons$, and 4) the growth rate in unemployment Δue. These measures coincide with those studied by Næs et al. (2011). Predictive regressions include two lags of the dependent variable for measures of real economic activity, and no lags for excess return-forecasting regressions. We consider predictive ability both in univariate models (in the sense of including only a single illiquidity measure) and in specifications that include a set of additional forecasting variables derived from the bond markets (the short-rate and various spread measures). Because the principal component is extracted using the full sample of data, we report results for break- and volatility-adjusted measures using the FS approach. This choice is not critical in terms of the pattern of predictability, however, and we obtain similar results under the RT approach.

Table 9 presents results. Several volatility- and break-adjusted illiquidity measures significantly forecast stock returns. In addition to illustrating robustness of monthly

Table 8
Monthly forecasts including macroeconomic and financial variables.

The table presents one-month-ahead predictive regressions for growth in industrial production (Δip) and growth in unemployment (Δue) using various measures of aggregate illiquidity:

$$y_{t+1} = \alpha + \sum_{j=1}^{J} \theta_j y_{t+1-j} + \beta x_t + \gamma' z_t + \epsilon_{t+1},$$

where y_{t+1} is the dependent variable and x_t indicates a break- and volatility-adjusted aggregate illiquidity measure of interest. We set $J = 4$ for both Δip and Δue. The vector z_t represents an additional set of macroeconomic and financial forecasting variables included in the model. We consider two choices for z_t. z_1 consists of lagged volatility (*vol*). z_2 consists of lagged volatility as well as the lagged term spread, default spread, commercial paper-to-Treasury spread, and stochastically detrended T-bill rate. For each forecasting variable, the table presents the estimated slope coefficient $\hat{\beta}$, the associated t-statistic based on Newey–West standard errors, and ΔR^2, defined as the increase in R^2-value relative to a similar regression that excludes x_t (the above model with $\beta = 0$). The sample period is 1948.3–2015.12 for Δip and 1948.6–2015.12 for Δue. See Section 2 of the paper or the Internet Appendix for variable definitions.

	$y_{t+1} = \Delta ip$						$y_{t+1} = \Delta ue$					
	Controls $= z_1$			Controls $= z_2$			Controls $= z_1$			Controls $= z_2$		
	$\hat{\beta}$	t-stat	ΔR^2	$\hat{\beta}$	t-stat	ΔR^2	$\hat{\beta}$	t-stat	ΔR^2	$\hat{\beta}$	t-stat	ΔR^2
Panel A: Break- and volatility-adjusted measures (RT approach)												
roll	0.01	0.05	0.00	−0.10	−0.44	0.01	0.95	1.84	0.27	1.94	3.57	1.04
cs	−0.01	−0.07	0.00	−0.08	−0.59	0.03	0.81	1.52	0.25	2.52	4.11	1.83
fht	0.11	0.62	0.03	0.08	0.45	0.01	0.61	1.26	0.13	0.96	1.96	0.30
tick	0.22	1.48	0.14	0.19	1.31	0.10	0.69	1.58	0.21	0.88	2.01	0.34
zeros	0.19	0.88	0.05	0.16	0.82	0.03	0.97	1.53	0.21	0.99	1.56	0.21
ami	−0.03	−0.53	0.01	−0.04	−0.68	0.02	0.46	2.39	0.55	0.64	3.28	1.00
amito	0.26	2.00	0.27	0.27	2.33	0.30	1.04	3.12	0.74	0.86	2.46	0.43
PS	0.28	0.19	0.00	0.69	0.46	0.01	2.71	0.73	0.03	−1.39	−0.38	0.01
hm	−0.25	−0.77	0.06	−0.11	−0.35	0.01	−0.53	−0.73	0.05	−1.22	−1.62	0.27
Panel B: Break- and volatility-adjusted measures (SD approach)												
roll	−0.03	−0.11	0.00	−0.01	−0.05	0.00	1.24	1.96	0.30	1.18	1.91	0.27
cs	−0.35	−0.84	0.13	−0.32	−0.76	0.10	3.59	2.46	1.06	3.75	2.44	1.03
fht	−0.14	−0.50	0.02	−0.14	−0.43	0.02	1.45	2.38	0.48	1.15	1.74	0.25
tick	−0.09	−0.28	0.01	−0.05	−0.15	0.00	1.92	2.92	0.73	1.68	2.33	0.46
zeros	−0.10	−0.37	0.01	−0.15	−0.47	0.02	1.22	1.84	0.29	0.97	1.31	0.15
ami	−0.19	−1.55	0.13	−0.15	−1.18	0.07	1.65	4.88	2.12	1.37	3.83	1.29
amito	−0.10	−0.48	0.02	−0.11	−0.53	0.02	1.49	2.96	0.68	1.45	2.71	0.58
PS	0.21	0.14	0.00	0.01	0.01	0.00	−0.34	−0.09	0.00	−1.83	−0.47	0.01
hm	−0.53	−1.16	0.17	−0.49	−1.04	0.14	−0.04	−0.04	0.00	−0.46	−0.54	0.02

results, it is notable that several less-successful illiquidity measures at the monthly horizon perform considerably better at the quarterly horizon. These include *roll*, *hm*, and *PS*. A potential explanation for performance improvements in quarterly data is that constructing measures at the longer quarterly frequency reduces the relative impact of noise. (Each of these measures relies on within-period parameter estimates obtained from underlying daily data.) The associated return predictability is economically significant, with R^2-values in excess of 2% for several predictors. Turning to measures of economic activity, aggregate volatility forecasts real activity for quarterly data (top line of Panel A). Most volatility-adjusted measures are insignificant in univariate specifications. As in monthly data, however, including additional controls (Panel B) strengthens the case that residual illiquidity measures contain predictive information for economic activity. In fact, once additional controls are included at the quarterly frequency, lagged volatility is in some cases insignificant, whereas many of the adjusted illiquidity measures become significant.

The principal component extracted from the underlying adjusted illiquidity measures (denoted *PC* in results) successfully predicts excess stock returns as well as most measures of economic activity (consumption growth is an exception). The *PC* variable is typically not the best-performing measure for particular forecasting targets. It performs relatively well for a variety of targets, however, and from this perspective appears to be a useful summary measure of (volatility-adjusted) illiquidity. Results reported in the Internet Appendix confirm that the principal component extracted at the monthly frequency significantly forecasts stock returns and unemployment growth; however, it is insignificant for growth in industrial production.

We conduct a variety of additional robustness checks concerning our main results. Stock return-forecasting regressions in the main paper employ the log excess return on the NYSE index as a market return proxy. However, evidence of stock return predictability is robust to the use of alternative market return measures such as the log excess return on the CRSP value-weighted universe or the log excess return the S&P 500 Index. Predictive regression results are qualitatively similar under alternative approaches to adjusting for volatility. For example, results for adjusted measures constructed based on regressions of illiquidity measures on log volatility (employing either OLS or narrow-band least squares slope estimates) produce similar results to those presented. Finally, we examine longer-horizon (up to one year) predictive regressions for excess returns and measures of economic activity. Granger causality test results for longer-horizon regressions are qualitatively similar to those at the monthly and

Table 9
Forecasting regressions: quarterly frequency.
The table presents results for one-quarter-ahead predictive regressions for excess returns on the NYSE index (*ret*), growth in GDP (Δgdp), growth in unemployment (Δue), growth in investment (Δinv), and growth in consumption ($\Delta cons$) using volatility (*vol*) and various break- and volatility-adjusted measures of aggregate illiquidity (FS approach). The table reports results from regressions of the form

$$y_{t+1} = \alpha + \sum_{j=1}^{J} \theta_j y_{t+1-j} + \beta x_t + \gamma' z_t + \epsilon_{t+1},$$

where y_{t+1} is the dependent variable, x_t denotes the forecasting variable of interest, and z_t captures additional predictors. *PC* denotes the first principal component extracted from the full set of underlying illiquidity measures. The table presents the estimated slope coefficient $\hat{\beta}$, the associated *t*-statistic based on Newey–West standard errors, and ΔR^2, defined as the increase in R^2-value relative to a similar regression that excludes x_t (the above model with $\beta = 0$). When the dependent variable is Δgdp, Δue, Δinv, or $\Delta cons$; two lags of the dependent variable are included in the predictive regression ($J = 2$). The predictive regression for excess returns does not include lagged returns. The sample period is 1948.1–2015.4 for *ret*, 1948.4–2015.4 for Δue, and 1948.2–2015.4 for other macroeconomic targets. Section 2 of the paper or the Internet Appendix gives variable definitions.

	$y_{t+1} = ret$			$y_{t+1} = \Delta gdp$			$y_{t+1} = \Delta ue$			$y_{t+1} = \Delta inv$			$y_{t+1} = \Delta cons$		
	$\hat{\beta}$	*t*-stat	ΔR^2	$\hat{\beta}$	*t*-stat	ΔR^2	$\hat{\beta}$	*t*-stat	ΔR^2	$\hat{\beta}$	*t*-stat	ΔR^2	$\hat{\beta}$	*t*-stat	ΔR^2
Panel A: Univariate regressions (z_t omitted)															
vol	−0.18	−0.07	0.00	−0.69	−2.86	2.92	3.34	2.01	1.31	−2.91	−2.17	2.23	−0.54	−2.47	2.39
roll	3.02	1.82	1.18	−0.16	−0.73	0.23	0.73	0.51	0.09	−1.16	−0.98	0.48	−0.09	−0.47	0.10
cs	1.14	0.63	0.16	0.17	0.76	0.24	−0.85	−0.53	0.11	0.61	0.49	0.13	0.20	0.97	0.43
fht	4.80	1.99	1.60	0.25	0.87	0.29	−1.89	−1.11	0.31	0.95	0.66	0.17	0.25	1.05	0.41
tick	4.00	2.26	1.66	0.02	0.07	0.00	−0.03	−0.02	0.00	−0.21	−0.17	0.01	0.05	0.25	0.03
zeros	6.68	2.25	1.84	0.14	0.39	0.06	−1.33	−0.56	0.09	0.25	0.12	0.01	0.23	0.73	0.19
ami	2.19	3.23	2.78	−0.06	−0.58	0.13	0.44	0.67	0.13	−0.48	−0.92	0.37	−0.02	−0.20	0.02
amito	1.51	0.64	0.26	−0.29	−1.34	0.64	2.73	2.07	1.02	−2.10	−1.76	1.36	−0.05	−0.26	0.03
PS	67.48	2.93	2.61	−4.97	−1.71	0.96	28.66	1.78	0.58	−25.10	−1.90	1.00	−4.05	−1.63	0.86
hm	−9.86	−5.48	4.91	0.65	2.05	1.45	−4.51	−1.70	1.27	4.47	2.14	2.80	0.44	1.57	0.90
PC	0.52	2.33	1.69	0.00	0.04	0.00	−0.01	−0.04	0.00	−0.04	−0.27	0.04	0.01	0.40	0.07
Panel B: $z_t = (ret_t, term_t, def_t, cpsp_t, tbl_t)'$															
vol	−0.84	−0.31	0.04	−0.52	−2.10	1.11	2.01	1.14	0.31	−1.66	−1.39	0.46	−0.54	−2.16	1.58
roll	3.55	1.92	1.32	−0.58	−2.45	2.29	3.32	2.29	1.40	−3.57	−2.90	3.55	−0.35	−1.75	1.07
cs	0.83	0.45	0.06	−0.50	−1.84	1.32	3.93	2.08	1.52	−3.39	−2.53	2.46	−0.18	−0.71	0.22
fht	4.67	1.97	1.22	−0.26	−0.88	0.24	2.64	1.46	0.47	−2.52	−1.75	0.94	−0.07	−0.26	0.02
tick	3.69	2.16	1.20	−0.45	−1.90	1.14	3.65	2.50	1.39	−2.99	−2.42	2.07	−0.29	−1.30	0.65
zeros	7.03	2.43	1.77	−0.33	−0.87	0.25	3.21	1.32	0.44	−3.26	−1.67	1.02	0.00	0.01	0.00
ami	2.17	3.23	2.44	−0.23	−2.27	1.72	1.66	2.67	1.65	−1.42	−2.69	2.72	−0.15	−1.48	0.96
amito	2.74	1.29	0.74	−0.35	−1.67	0.76	2.57	2.32	0.74	−2.20	−1.98	1.26	0.00	0.02	0.00
PS	93.80	3.50	4.48	−1.01	−0.35	0.04	0.49	0.03	0.00	−6.38	−0.49	0.06	−0.07	−0.03	0.00
hm	−12.65	−6.00	7.09	0.32	1.01	0.31	−2.20	−0.79	0.27	3.12	1.42	1.20	0.09	0.29	0.03
PC	0.53	2.54	1.50	−0.06	−1.90	1.33	0.48	2.28	1.42	−0.44	−2.51	2.60	−0.03	−0.97	0.38

quarterly horizons. The economic significance of the predictive power afforded by the illiquidity measures tends to increase with the forecast horizon. This is particularly true for stock return forecasts. Slope coefficients and R^2-values increase noticeably for 6- and 12-month return forecasts relative to one-month-ahead results. These horizon effects are similar to long-run stock return forecasts for other persistent predictors in the literature, most notably financial ratios such as the dividend and earnings yield.

5. Conclusion

This paper constructs and analyzes long time series of measures of illiquidity. Illiquidity measures contain a component related to the US business cycle. We decompose illiquidity proxies into a component capturing aggregate volatility and a residual component. We also adjust for structural shifts driven by tick-size reductions. Empirically, volatility- and break-adjustments clarify the predictive content in aggregate illiquidity measures. Specifically, we find strong and robust evidence that volatility- and break-adjusted illiquidity measures forecast stock market returns. In addition, we find that both the volatility and residual components of illiquidity contain information regarding future real economic activity.

The illiquidity measures we construct potentially constitute the basis for new empirical work on liquidity and asset prices. An important strand of literature considers liquidity to be a priced risk factor (see, e.g., Pástor and Stambaugh, 2003; Acharya and Pedersen, 2005). Existing empirical literature tests whether securities' exposure to shocks to aggregate illiquidity (measured by proxies) is priced. This paper shows that most aggregate illiquidity proxies contain a component reflecting aggregate volatility, and that the predictive content associated with this component differs from the residual component. These results motivate separately testing for cross-sectional variation in expected returns as a function of exposure to shocks to each component.

The structure of equity markets continues to evolve, along with data available to researchers interested in measuring liquidity. The availability of high-frequency, intraday data in recent decades affords the potential to compute improved liquidity proxies. At the same time, ongoing changes in equity market structure present new challenges for empirical researchers. For example, 'tick-size wars' among competing exchanges (Meling and Ødegaard, 2016) and the increasing prevalence of alternative trading systems potentially alter the relative informativeness of various illiquidity proxies. Consequently, there remains much scope for future research regarding the measurement of aggregate liquidity conditions.

References

Acharya, V.V., Pedersen, L.H., 2005. Asset pricing with liquidity risk. J. Financial Econ. 77 (2), 375–410.
Ahn, H.-J., Cao, C.Q., Choe, H., 1998. Decimalization and competition among stock markets: evidence from the Toronto Stock Exchange cross-listed securities. J. Financial Mark. 1 (1), 51–87.
Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. J. Financial Mark. 5 (1), 31–56.
Andersen, T.G., Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial markets. J. Empir. Finance 4 (2), 115–158.
Bacidore, J., Battalio, R.H., Jennings, R.H., 2003. Order submission strategies, liquidity supply, and trading in pennies on the New York Stock Exchange. J. Financial Mark. 6 (3), 337–362.
Bacidore, J.M., 1997. The impact of decimalization on market quality: an empirical investigation of the Toronto Stock Exchange. J. Financial Intermed. 6 (2), 92–120.
Baillie, R.T., Bollerslev, T., Mikkelsen, H.O., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. J. Economet. 74 (1), 3–30.
Bansal, R., Kiku, D., Shaliastovich, I., Yaron, A., 2014. Volatility, the macroeconomy, and asset prices. J. Finance 69 (6), 2471–2511.
Beber, A., Brandt, M.W., Kavajecz, K.A., 2011. What does equity sector orderflow tell us about the economy? Rev. Financial Stud. 24 (11), 3688–3730.
Bekaert, G., Harvey, C.R., Lundblad, C., 2007. Liquidity and expected returns: Lessons from emerging markets. Rev. Financial Stud. 20 (6), 1783–1831.
Beran, J., 1995. Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. J. R. Stat. Soc. Ser. B 57, 659–672.
Bessembinder, H., 2003. Trade execution costs and market quality after decimalization. J. Financial Quant. Anal. 38 (04), 747–777.
Boehmer, E., Saar, G., Yu, L., 2005. Lifting the veil: an analysis of pre-trade transparency at the NYSE. J. Finance 60 (2), 783–815.
Brennan, M., Huh, S.W., Subrahmanyam, A., 2013. An analysis of the Amihud illiquidity premium. Rev. Asset Pric. Stud. 3 (1), 133–176.
Brockman, P., Chung, D.Y., Pérignon, C., 2009. Commonality in liquidity: A global perspective. J. Financial Quant. Anal. 44 (04), 851–882.
Brunnermeier, M.K., Pedersen, L.H., 2009. Market liquidity and funding liquidity. Rev. Financial Stud. 22 (6), 2201–2238.
Campbell, J.Y., 1991. A variance decomposition for stock returns. Econ. J. 101 (405), 157–179.
Campbell, J.Y., Lettau, M., Malkiel, B.G., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. J. Finance 56 (1), 1–43.
Campbell, J.Y., Thompson, S.B., 2008. Predicting excess stock returns out of sample: can anything beat the historial average? Rev. Financial Stud. 21, 1509–1531.
Campbell, J.Y., Yogo, M., 2006. Efficient tests of stock return predictability. J. Financial Econ. 81 (1), 27–60.
Cao, C., Chen, Y., Liang, B., Lo, A.W., 2013. Can hedge funds time market liquidity? J. Financial Econ. 109 (2), 493–516.
Chakravarty, S., Wood, R.A., Van Ness, R.A., 2004. Decimals and liquidity: a study of the NYSE. J. Financial Res. 27 (1), 75–94.
Chen, S.S., Chou, Y.H., Yen, C.Y., 2016. Predicting US recessions with stock market illiquidity. BE J. Macroecon. 16 (1), 93–123.
Chordia, T., Roll, R., Subrahmanyam, A., 2000. Commonality in liquidity. J. Financial Econ. 56 (1), 3–28.
Chordia, T., Roll, R., Subrahmanyam, A., 2001. Market liquidity and trading activity. J. Finance 56 (2), 501–530.
Chordia, T., Sarkar, A., Subrahmanyam, A., 2005. An empirical analysis of stock and bond market liquidity. Rev. Financial Stud. 18 (1), 85–129.
Christensen, B.J., Nielsen, M.Ø., 2006. Asymptotic normality of narrow-band least squares in the stationary fractional cointegration model and volatility forecasting. J. Economet. 133 (1), 343–371.
Clark, T.E., McCracken, M.W., 2011. Testing for unconditional predictive ability. In: Clements, M.P., Hendry, D.F. (Eds.), The Oxford Handbook of Economic Forecasting, vol. 1. Oxford University Press, New York, pp. 415–440.
Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. J. Economet. 138, 291–311.
Corwin, S.A., Schultz, P., 2012. A simple way to estimate bid-ask spreads from daily high and low prices. J. Finance 67 (2), 719–760.
Diebold, F.X., 2015. Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests. J. Bus. Econ. Stat. 33 (1), 1–24.
Ding, Z., Granger, C.W., Engle, R.F., 1993. A long memory property of stock market returns and a new model. J. Empir. Finance 1 (1), 83–106.
Ferson, W.E., Sarkissian, S., Simin, T.T., 2003. Spurious regressions in financial economics? J. Finance 58 (4), 1393–1413.
Florackis, C., Giorgioni, G., Kostakis, A., Milas, C., 2014. On stock market illiquidity and real-time GDP growth. J. Int. Money Finance 44, 210–229.
Fong, K., Holden, C., Trzcinka, C., 2017. What are the best liquidity proxies for global research? Rev. Finance 21, 1355–1401.
Fornari, F., Mele, A., 2013. Financial volatility and economic activity. J. Financial Manag. Mark. Inst. 1 (2), 155–198.
French, K.R., Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. J. Financial Econ. 19 (1), 3–29.
Ghysels, E., Santa-Clara, P., Valkanov, R., 2005. There is a risk-return trade-off after all. J. Financial Econ. 76 (3), 509–548.
Giacomini, R., White, H., 2006. Tests of conditional predictive ability. Econometrica 74, 1545–1578.
Giglio, S., Kelly, B.T., Pruitt, S., 2016. Systemic risk and the macroeconomy: an empirical evaluation. J. Financial Econ. 119, 457–471.
Goldstein, M.A., Kavajecz, K.A., 2000. Eighths, sixteenths, and market depth: changes in tick size and liquidity provision on the NYSE. J. Financial Econ. 56 (1), 125–149.
Goyal, A., Welch, I., 2008. A comprehensive look at the empirical performance of the equity premium prediction. Rev. Financial Stud. 21, 1455–1508.
Goyenko, R.Y., Holden, C.W., Trzcinka, C.A., 2009. Do liquidity measures measure liquidity? J. Financial Econ. 92 (2), 153–181.
Granger, C.W., Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. J. Time Ser. Anal. 1 (1), 15–29.
Harris, L.E., 1994. Minimum price variations, discrete bid–ask spreads, and quotation sizes. Rev. Financial Stud. 7 (1), 149–178.
Harris, L.E., 1997. Decimalization: A Review of the Arguments and Evidence. Unpublished working paper. University of Southern California.
Hasbrouck, J., Seppi, D.J., 2001. Common factors in prices, order flows, and liquidity. J. Financial Econ. 59 (3), 383–411.
Hodrick, R.J., 1992. Dividend yields and expected stock returns: alternative procedures for inference and measurement. Rev. Financial Stud. 5 (3), 357–386.
Holden, C.W., 2009. New low-frequency spread measures. J. Financial Mark. 12 (4), 778–813.
Hosking, J.R., 1981. Fractional differencing. Biometrika 68 (1), 165–176.
Hou, K., Moskowitz, T.J., 2005. Market frictions, price delay, and the cross–section of expected returns. Rev. Financial Stud. 18 (3), 981–1020.
Huberman, G., Halka, D., 2001. Systematic liquidity. J. Financial Res. 24 (2), 161–178.
Jansson, M., Moreira, M.J., 2006. Optimal inference in regression models with nearly integrated regressors. Econometrica 74 (3), 681–714.
Jones, C.M., 2002. A Century of Stock Market Liquidity and Trading Costs. Unpublished working paper. Columbia University.
Jones, C.M., Lipson, M.L., 2001. Sixteenths: direct evidence on institutional execution costs. J. Financial Econ. 59 (2), 253–278.
Karolyi, G.A., Lee, K.-H., Van Dijk, M.A., 2012. Understanding commonality in liquidity around the world. J. Financial Econ. 105 (1), 82–112.
Kayacetin, N.V., Kaul, A., 2009. Forecasting Economic Fundamentals and Stock Returns with Equity Market Order Flows. Unpublished working paper. University of Alberta.
Killian, L., 1999. Exchange rates and monetary fundamentals: What do we learn from long-horizon regressions? J. Appl. Economet. 14 (5), 491–510.
Korajczyk, R.A., Sadka, R., 2008. Pricing the commonality across alternative measures of liquidity. J. Financial Econ. 87 (1), 45–72.
Kyle, A.S., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.
Lesmond, D.A., Ogden, J.P., Trzcinka, C.A., 1999. A new estimate of transaction costs. Rev. Financial Stud. 12 (5), 1113–1141.
Lettau, M., Van Nieuwerburgh, S., 2008. Reconciling the return predictability evidence. Rev. Financial Stud. 21 (4), 1607–1652.
Mancini, L., Ranaldo, A., Wrampelmeyer, J., 2013. Liquidity in the foreign exchange market: measurement, commonality, and risk premiums. J. Finance 68 (5), 1805–1841.
Mark, N.C., 1995. Exchange rates and fundamentals: Evidence on long-horizon predictability. Am. Econ. Rev. 85 (1), 201–218.
Maynard, A., Phillips, P.C., 2001. Rethinking an old empirical puzzle: econometric evidence on the forward discount anomaly. J. Appl. Economet. 16 (6), 671–708.

Maynard, A., Smallwood, A., Wohar, M.E., 2013. Long memory regressors and predictive testing: a two-stage rebalancing approach. Economet. Rev. 32 (3), 318–360.
Meling, T.G., Ødegaard, B.A., 2016. Tick Size Wars. Unpublished working paper. University of Bergen.
Merton, R.C., 1973. An intertemporal capital asset pricing model. Econometrica 41, 867–887.
Næs, R., Skjeltorp, J.A., Ødegaard, B.A., 2011. Stock market liquidity and the business cycle. J. Finance 66 (1), 139–176.
Pástor, L., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. J. Polit. Econ. 111 (3), 642–685.
Rai, A., 2015. Stock market illiquidity's predictive role over economic growth: the Australian evidence. Rev. Pacific Basin Financial Mark. Policies 18 (02), 1–30.
Rapach, D.E., Strauss, J.K., Zhou, G., 2010. Out-of-sample equity premium prediction: combination forecasts and links to the real economy. Rev. Financial Stud. 23, 821–862.
Robinson, P.M., 1994. Semiparametric analysis of long-memory time series. Ann. Stat. 22, 515–539.
Robinson, P.M., Marinucci, D., 2003. Semiparametric frequency domain analysis of fractional cointegration. In: Robinson, P.M. (Ed.), Time Series with Long Memory, Vol. 1. Oxford University Press, New York, pp. 334–373.
Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. J. Finance 39 (4), 1127–1139.
Seppi, D.J., 1997. Liquidity provision with limit orders and a strategic specialist. Rev. Financial Stud. 10 (1), 103–150.
Shimotsu, K., Phillips, P.C., 2005. Exact local Whittle estimation of fractional integration. Ann. Stat. 33 (4), 1890–1933.
Smimou, K., 2014. Consumer attitudes, stock market liquidity, and the macro economy: A Canadian perspective. Int. Rev. Financial Anal. 33, 186–209.
Smimou, K., Khallouli, W., 2015. Does the Euro affect the dynamic relation between stock market liquidity and the business cycle? Emerg. Mark. Rev. 25, 125–153.
Stambaugh, R.F., 1999. Predictive regressions. J. Financial Econ. 54 (3), 375–421.
Subrahmanyam, A., 2007. Liquidity, return and order-flow linkages between REITs and the stock market. Real Estate Econ. 35 (3), 383–408.
Van Ness, B.F., Van Ness, R.A., Pruitt, S.W., 2000. The impact of the reduction in tick increments in major US markets on spreads, depth, and volatility. Rev. Quant. Finance Account. 15 (2), 153–167.
West, K.D., 1996. Asymptotic inference about predictive ability. Econometrica 64 (5), 1067–1084.