Noname manuscript No.
(will be inserted by the editor)

Run Lengths and Liquidity

December 14, 2008

Abstract We develop a market-wide illiquidity risk factor based on run lengths and find that it is priced using standard asset-pricing specifications. Our theoretical framework of equity returns derives the result that average run lengths of individual stocks proxy for illiquidity, and are related to common measures of liquidity such as trading volume and trade price-impact. This relationship holds irrespective of the sampling frequency in the computation of run lengths. Thus, liquidity can be quantified by examining a stock's run length signature, providing a statistical mechanics link across illiquidity metrics. Tests using daily equity return data for all stocks over the period 1962-2005 find that run lengths are decreasing in turnover, and increasing with bid-ask spreads, and price-impact. Illiquidity is shown to be a risk factor/characteristic in explaining equity returns.

Keywords run length · liquidity
JEL codes: G0, G1, G12

Thanks to Viral Acharya, George Chacko, Tarun Chordia, Jennifer Huang, Michael Pagano, George Penacchi, Priya Raghubir, Atulya Sarin, Mark Seasholes, Marti Subrahmanyam and seminar participants at Villanova University, Calpers, QWAFAFEW San Francisco, FDIC, and the University of Chicago conference on liquidity for helpful discussions. We are especially grateful to Yakov Amihud and an unknown referee for numerous suggestions and ideas. The first author acknowledges financial support from the Dean Witter Foundation and a Breetwor Fellowship. Contact the authors at Sanjiv Das, Santa Clara University, `srdas@scu.edu` and Paul Hanouna, Villanova University, `paul.hanouna@villanova.edu`.

Sanjiv R. Das
Santa Clara University, Leavey School of Business,
500 El Camino Real, Santa Clara, California, 95053, USA.
`srdas@scu.edu`

Paul Hanouna
Villanova University, Villanova School of Business,
800 Lancaster Avenue, Villanova, Pennsylvania, 19085, USA.
`paul.hanouna@villanova.edu`

1 Introduction

This paper presents run length as a measure of liquidity in equity markets. This metric is simple to compute, and tracks liquidity very well. Average run length is inversely related to trading intensity[1], and is positively related to the price impact of a trade. These are two commonly used measures of liquidity in both, theoretical and empirical work.[2]

Runs have been examined in the context of efficiency, over-reaction, and serial dependence, but not liquidity. Specifically, runs have previously been used to examine the informational efficiency of stocks, since a random walk leaves a distinctive run length signature.[3] Fama (1965) computed both, the number of runs and the length of runs for several stocks, and rejected violations of efficiency based on serial dependence in returns. Easley, Kiefer, and O'Hara (1997) used runs to examine dependence in intra-day data to estimate a dynamic model of market-maker behavior. McQueen and Thorley (1984) use runs to analyze stock market bubbles. In this paper, we demonstrate theoretically and empirically that runs are a good proxy for stock liquidity. The theoretical model shows that illiquid stocks (evidencing low turnover, large price impact) have longer run lengths. Asset pricing tests show that a liquidity factor constructed from run length sorted portfolios explains the cross section of returns, even after controlling for the Fama-French and momentum asset pricing factors.

Liquidity has been defined in myriad ways, and is gaining traction as a pertinent factor in asset pricing [see Amihud, Mendelson and Pedersen (2005) for a wide-ranging survey of the literature on this topic, and Wyss (2004) for a comprehensive list of liquidity measures]. Recent work has introduced different metrics of liquidity, such as the illiquidity measure of Amihud (2002), that measures price impact per dollar trading volume. Pastor and Stambaugh (2003) develop a measure based on order flow induced reversals. Liu (2003) develops a measure of liquidity based on no-trade days. Chacko (2004) builds a bond market liquidity measure that is based on the proportion of total stock in a bond that is available for trading. Chacko, Jurek and Stafford (2008) develop a metric of equity market liquidity based on the value of an option to trade out of a position in the stock, and Sadka (2006)'s measure is based on order flow and price impact. Korajczyk and Sadka (2006) examine whether these varied metrics essentially measure the same phenomenon, and find that there is a common component across the many different definitions of liquidity. Our run-based liquidity measure is parsimonious and may be computed using only a stock's price series. In contrast, other measures require trading volumes, order flow, price impact, and inventory data as well. Hence, the run length measure may be applied in international markets where data is not as ubiquitous as in the U.S.[4]

It is expensive to trade in illiquid stocks. Bid-ask spreads tend to be larger. Hence, illiquid stocks will be characterized by lower trade arrival rates and higher price impact. Trading faster and in larger quantities in liquid stocks than in illiquid ones, results in incorporating information faster into prices of liquid stocks, and slower for illiquid ones.[5] This effect is responsible for the evidence that post-earnings-announcement drift

[1] See Pagano (1989) for a model of concentration of trading volume.

[2] See Brennan and Subrahmanyam (1996) for premia related to price-impact.

[3] Pure random walks (e.g. unbiased coin tosses) have an average run length of 2.

[4] We are grateful to Yakov Amihud for pointing this out to us.

[5] Thanks to the referee for this real-world explanation connecting our physical and financial intuition.

occurs mainly for highly illiquid stocks (see Chordia, Goyal, Sadka, Sadka and Shivakumar (2006)). Our run-length metric is consistent with the phenomenon that higher costs of trading will result in lower deal flow, persistent price impact and thus, longer run lengths.

The classic paper by Amihud and Mendelson (1986) shows that a widely used proxy for liquidity, the bid-ask spread, is related systematically to returns on stocks. We show that bid-ask spreads are positively related to the run length measure.[6] We also show that run length explains stock returns in the cross-section.[7] This complements the growing literature that finds compensation for the providers of liquidity [see for example Pastor and Stambaugh (2003), Liu (2003), Acharya and Pedersen (2005), Sadka (2006), Korajczyk and Sadka (2006), and Li, Mooradian and Zhang (2007)]. Thus, the run length measure is not only easy to compute, but also has validity for asset pricing. The measure may also be used as a characteristic in pricing regressions, and we show that it is priced even after including other known firm characteristics that are used in this literature (see Daniel and Titman (1997); Amramov and Chordia (2001)).

It is possible to reinterpret much of the extant literature on stock return persistence, autocorrelation, reversals and momentum in terms of runs. For instance, our theoretical measure and empirics are consistent with the findings of Conrad, Hameed, and Niden (1994) that high transaction (or high volume) securities evidence greater reversals (lower run lengths). The signed trading volume measure of illiquidity developed by Pastor and Stambaugh (2003) is consistent with the logic that increases in order flow result in shorter runs, as market-makers earn premia from injecting liquidity into the market. Campbell, Grossman and Wang (1993) find that increased trading volume reduces serial dependence, analogous to shorter runs. Hendershott and Seasholes (2007) show that liquidity injections through inventory build-ups by market makers is followed by reversals (shorter runs) that make it profitable for them to act as liquidity providers. Run lengths are positively correlated with the zero-return day count measure introduced by Lesmond, Ogden and Trzcinka (1999) to proxy for transaction costs, and with the no-trade day count liquidity measure of Liu (2003). Avramov, Chordia and Goyal (2006) show that liquidity infusions, after controlling for trading volume, result in price reversals (shorter runs), and Chordia, Roll and Subrahmanyam (2005) provide evidence that liquidity reduces autocorrelation in returns (shorter runs). In contrast, Getmansky, Lo and Makarov (2004) show that illiquidity induces serial correlation in hedge fund returns and is closer in spirit to the model in this paper. Finally, traders often break up a large trade into a series of smaller trades when liquidity is low, and this would induce longer run lengths. These connections offer intuition for the mechanics of the relation between liquidity and runs. Effects such as these were in evidence in the recent sharp market decline on Tuesday, February 27, 2007, (known as "Grey Tuesday"), where prices moved monotonically downward, and were subsequently reversed by liquidity injections.

One might intuitively expect that run length and momentum are similar. But, there are two differences. First, run length is a short run phenomenon, and reversals in price series are frequent, whereas momentum is usually measured and assessed for returns

[6] Our theoretical result that stocks with shorter runs have lower price impact is consistent with the research in Hasbrouck (1991), Brennan and Subrahmanyam (1996) and Hasbrouck (2006) relating bid-ask spreads to premia and information content.

[7] Liquidity is important not only in primary equity markets, but also in derivatives markets [see Cetin, Jarrow, Protter and Warachka (2006)].

over much longer horizons as in the literature spawned by Jegadeesh and Titman (1993). Second, returns to momentum come from holding winners and shorting losers (as per the measure of Carhart (1997)), where the sign of the persistence (long term run) matters. As runs gets longer (irrespective of sign), our model suggests that illiquidity increases and providers of liquidity will earn premia. Further, if our run length measure were purely a proxy for momentum, then after controlling for it, runs should not explain the cross-section of returns in asset-pricing tests. Our measure's explanatory power persists despite momentum controls, indicating that runs are distinct liquidity metrics and not momentum related.

Briefly, here are the main results of this paper.

1. Whereas other work on illiquidity uses variables that proxy for the price impact of a trade, our run-length metric uncovers the statistical mechanics of liquidity. Run length may be viewed as a sufficient statistic for trading volume and price impact. As we will see, it is mathematically connected to liquidity via the stochastic process of the traded security.
2. We construct an illiquidity factor using run length based portfolios, and show that, using standard asset pricing tests, this factor is priced over a large sample period.
3. Our measure of liquidity is also shown to be a priced characteristic that remains statistically significant after controlling for other firm characteristics used in the literature.
4. We empirically examine the connection between run length and other liquidity proxies using more than forty years of data and find a strong relationship. The run length measure of liquidity is shown to reside within the space spanned by other widely used measures. It offers a parsimonious alternative to other metrics in the literature.

The paper proceeds as follows. In Section 2, we derive the new metric of liquidity. Section 3 describes the data: all stocks in the CRSP database from 1962-2005. Section 4 shows that trading volume is inversely related to run length, and positively related to the price impact of a trade. Hence, the metric sorts firms correctly, no matter whether we consider liquidity in its trading volume connotation or as the price impact of trades. Section 5 consolidates our findings by demonstrating that our new liquidity measure explains both bid-ask spreads as well as the Amihud (2002) measure of illiquidity, even after controlling for trading volume and auto-correlation in returns. In Section 6 we construct a factor-mimicking portfolio for liquidity and show that it performs well in asset-pricing tests, demonstrating that liquidity is priced, even after controlling for the Fama-French factors and momentum. We also show that run length is a price firm characteristic. The section contains a principal components analysis showing how the run length measure relates to other measures of liquidity. Section 7 concludes.

2 Analysis

A run is a consecutive series of price moves without a sign reversal. Runs are simple constructs, yet little recent research has been devoted to them in finance, though there is a vast statistical literature from decades ago on the subject. (See Edgington (1961) for a useful statistic for counting series of positive and negative numbers, i.e. the number of runs.) Runs have been used to diagnose whether a series is a random walk

[Fama (1965), Moore (1978) and Grafton (1981)], useful in tests of the efficient market hypothesis. Here we are primarily interested in the average run length of a stock series.

We define h to be the inter-arrival time between trades. We analyze an i.i.d. process for stock return innovations, conditional on h, i.e.

$$R|h = \mu h + \sigma\epsilon\sqrt{h}, \quad \epsilon \sim N(0,1), \tag{1}$$

which implies a return per trade interval with mean μh and variance $\sigma^2 h$. The variables μ and σ are the expected return and standard deviation of return (per chosen unit time interval) respectively. The inter-arrival trade time h may be a function of market conditions and trading costs, and if illiquidity makes trading more expensive, h will increase.

The probability of a negative return for trade interval h is

$$p(h) = \Pr[R < 0] = \Pr[\mu h + \sigma\epsilon\sqrt{h} < 0] = \Phi\left(\frac{-\mu\sqrt{h}}{\sigma}\right) \tag{2}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. Note here that as $h \to 0$, this probability tends to one-half, because the variance swamps the mean.

Conditional on being in a positive run, the average length of such runs will depend on the probability that a negative return does *not* occur. To begin, say we are interested in the survival probability of a positive run. Define L^+ as the length of a positive run. Then,

$$\Pr[L^+ = n] = [1 - p(h)]^{n-1}\, p(h). \tag{3}$$

For example, if $n = 1$, the run is comprised of one positive return, and is terminated. The probability of this happening is $p(h)$.

The average run length for positive runs is then given by the following infinite sum:

$$\bar{L}^+ = \sum_{n=1}^{\infty} \left\{ n \times \Pr[L^+ = n] \right\} \tag{4}$$

$$= \sum_{n=1}^{\infty} \left\{ n \times [1 - p(h)]^{n-1}\, p(h) \right\} = \frac{1}{p(h)}. \tag{5}$$

By analogy, the average length of a negative run is

$$\bar{L}^- = \frac{1}{1 - p(h)}. \tag{6}$$

In the limit, since there are an equal number of positive and negative runs (over a long sequence of trades), the average run length will be the average of the averages of positive and negative run lengths, i.e.,

$$\begin{aligned}
\bar{L} &= \frac{1}{2}[\bar{L}^+ + \bar{L}^-] \\
&= \frac{1}{2}\left[\frac{1}{p(h)} + \frac{1}{1 - p(h)}\right], \\
&= \frac{1}{2}\left[\frac{1}{\Phi\left(\frac{-\mu\sqrt{h}}{\sigma}\right)} + \frac{1}{\Phi\left(\frac{\mu\sqrt{h}}{\sigma}\right)}\right]
\end{aligned} \tag{7}$$

where the second line follows from equations (5) and (6). The third line comes from the result in equation (2).

Example: if $\mu = 0$, then we have a symmetric random walk, and $p(h) = 0.5, \forall h$. In this case, $\bar{L} = 2$.

We note the following two results:

- *Result 1: Trading activity.* Run length is inversely related to trading activity (shorter trade inter-arrival times h), as may be seen from equation (7), where $\frac{\partial \bar{L}}{\partial h} > 0$. This embodies the simple idea that when there is high liquidity, trading activity rises, which implies that the time interval between trades (h) shrinks. When $h \to 0$, i.e. as the frequency of sampling increases, the average run length tends to 2.
  Corollary (Volatility): A known empirical regularity is that trading volume and volatility (σ) are positively related, see Karpoff (1987). Therefore, if volatility (related to trading volume) increases, equation (7) would indicate that the average run length would tend to decline ($\frac{\partial \bar{L}}{\partial \sigma} < 0$), and in the limit would be 2.
- *Result 2: Price Impact.* Trade price impact [see Glosten (1989), Chacko, Jurek and Stafford (2008) for models of this] in equation (1) is a function of the absolute return per unit risk (the ratio $|\mu|/\sigma$ in equation 7) from each trade (keeping h fixed). The average run length $\bar{L}$ increases in $|\mu|/\sigma$. This is intuitive because the tendency to drift in one direction increases. Hence, increasing average run length corresponds to greater price impact from each trade, or lower liquidity. Jones, Kaul and Lipson (1994) demonstrate that increases in trade arrivals (not generically trading volume) result in higher volatility and account for much of the price impact. In our framework, their result would imply that increasing trade arrivals will result in increases in σ and a reduction in average run lengths.

Note that *both* results indicate that *illiquidity* is related to longer run lengths. The simple intuition is that higher transaction arrival results in the variance of the process swamping its mean, reducing price impact, and simultaneously shortening average run length. We illustrate both results in Figure 1, which plots the average run length of the formula in equation (7) against (a) time interval h, and (b) drift μ (keeping σ fixed). When h tends to zero, the run length converges to 2, that of the symmetric random walk. The trading volume and price impact aspects of liquidity may be viewed as different sides of the Grossman and Miller (1988) characterization of liquidity being related to the price of immediacy. The cheaper immediacy is, the more frequent trading will be, and the price impact of trading will be lower (see Chacko, Jurek and Stafford (2008)).

If immediacy is in short supply, bid-ask spreads will widen, resulting in even lower immediacy, no trading, and hence zero returns. Therefore, long periods of flat prices are symptomatic of low liquidity and concurrently result in zero-return days and of course, long unbroken runs. We show that the zero-return measure is complementary to the run length metric. In the ensuing empirical sections this connection between trading intensity and runs will become clear.

Random trade arrivals and sampling

When transaction costs (bid-ask spreads) are high, liquidity in the market is dampened, and trade arrivals become infrequent, as pointed out in Lesmond, Ogden and Trzcinka (1999) and Liu (2003). In order to accommodate this source of variation, we extend the run length model to random trade arrivals. We note that the trade arrival rate may be

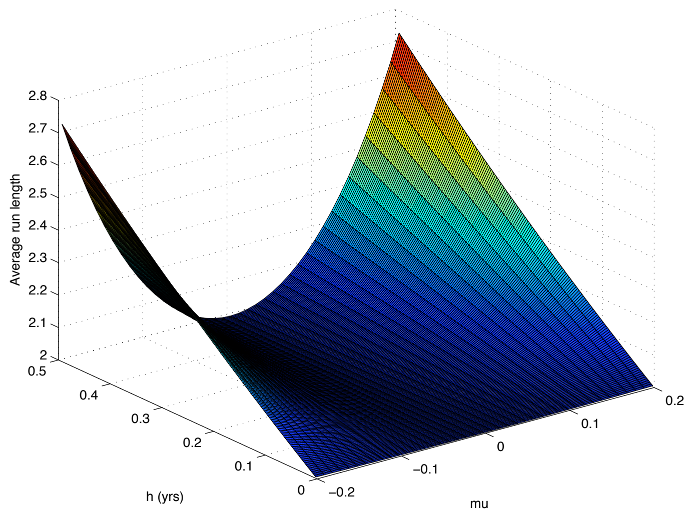

Fig. 1 Variation in average run length with changing μ and h. ($\sigma = 0.2$). Think of h as a proxy for transaction volume, and μ as a proxy for price impact.

modeled as a function of trading costs – as illiquidity makes trading more expensive trading frequency declines. As the inter-arrival time between trades increases, a strategy of sampling prices at fixed intervals will result in longer average run lengths. In this setting, zero returns/no trades becomes a component of the run length measure.

Suppose h depends on a random arrival of trades. If the rate at which trades arrive is Poisson with parameter $\lambda \equiv \lambda(c)$, where c is the cost of illiquidity, then the average inter-arrival time between trades will be $1/\lambda$ and will be positively correlated with trading costs c. The distribution of inter-arrival times will be exponential and the probability density of h will be $(1/\lambda)\exp[-h/\lambda]$. Simple extension of the previous calculations shows that the average run length will be given by

$$\bar{L} = \frac{1}{2}[\bar{L}^+ + \bar{L}^-] = \frac{1}{2}\left[\frac{1}{p(\lambda)} + \frac{1}{1-p(\lambda)}\right], \tag{8}$$

$$p(\lambda) = \int_0^\infty \Phi\left(\frac{-\mu\sqrt{h}}{\sigma}\right) \frac{1}{\lambda} e^{-h/\lambda}\, dh. \tag{9}$$

The results have now been expressed as a function of the trade arrival rate λ as opposed to the trade interval h. It is easily checked that Results 1 and 2 above are unaltered for random arrivals. As λ increases, the average trade inter-arrival time decreases, and the run length falls as well. Therefore, random arrivals of trades do not impact the results we obtained from constant trade inter-arrival times. Since $\frac{d\lambda}{dc} < 0$, increases in costs of illiquidity result in lower deal flow, and longer run lengths. Finally, we note that the average number of runs per unit time will be $\lambda/\bar{L}$.

Intuitively, these results should apply no matter what time units we use, since the results are not specific to the unit of time being a minute, hour or day. Also, if we sampled the stock price series at varied intervals, not just at each trade, the ordering of stocks by price run length is not affected on average, and in fact, as the sampling interval increases, the role of $|\mu|$ is enhanced, resulting in sharper differences amongst

Table 1 The effect of sampling interval on run lengths. We present the results of a simulation experiment to show that changes in sampling frequency do not change the ordering of stocks by run length. We assumed that trades arrive at Poisson frequency with mean λ trades per day for a total of 2600 days. We chose four different values of the trade arrival rate, $\lambda = \{0.25, 1, 10, 40\}$ trades per day, representing increasing frequency of trades (resulting in essentially four different stocks). We set $\mu = 0.1$ and $\sigma = 0.2$. We simulated price paths for all these stocks by generating inter-trade times (h) from an exponential distribution with parameter $1/\lambda$, and then returns using equation (1). Once the price series for all 2600 days is generated, we then sample the series at four different frequencies: (i) each trade, (ii) hourly, (iii) daily, (iv) weekly. The ordering of stocks by run length does not change with sampling frequency - note that the ordering of run length in each column has remained the same, even though the stock price is being sample less frequently as we go from left to right in the table. In fact, sampling less frequently exaggerates the difference in run lengths across the various trade frequencies. The numbers in columns 2 through 5 in the table below are the average run lengths for four stocks of different trade arrival rates (rows), sampled at four different frequencies (columns).

λ	Sampling frequency			
	Each trade	Hourly	Daily	Weekly
0.25	3.1000	3.4000	3.6122	12.0833
1.00	2.3648	2.2932	2.6023	5.3718
10.00	2.0167	2.0891	2.4351	4.7843
40.00	2.0100	2.0200	2.4235	4.2667

the run lengths, leaving the ordering intact. To exemplify this intuition, we undertook a simulation which is presented (details and results) in Table 1. It shows that the ordering of run lengths for stocks with varied trade intervals is unchanged when the sampling interval for prices is varied from trade-by-trade to hourly to daily and finally weekly. We also see that with a greater sampling interval the run length ordering appears more starkly.

Before proceeding on to the empirical analyses, we summarize the basic features of our runs specification: (a) Run lengths increase with the price impact of trades. (b) Run lengths decrease with trading volume. (c) Run length increases as trade inter-arrival time ($1/\lambda$) increases and is positively correlated with no-trade/zero-return days.

In the next sections, we present empirical analyses based on our run length measure: (a) We describe the daily data we use to analyze all stocks in the CRSP database for run lengths. (b) We demonstrate that run lengths are indeed related to trading volume. (c) We show that run length is related closely to the Amihud (2002) price-impact measure of liquidity. (d) We construct a liquidity factor, (e) we show that the liquidity factor explains the cross-section of stock returns even after controlling for other standard asset-pricing factors, (f) run length is shown to be a priced characteristic, and (g) a principal components analysis shows that it resides within the factor space of other widely used liquidity measures.

3 Data

Using the CRSP daily files we compute the yearly run length characteristics of each eligible stock on the NYSE, Amex, and Nasdaq exchanges from January 1962 to December 2005. Eligible stocks are common stocks with a year-end stock price between \$5 and \$1000, trading on the NYSE, AMEX and Nasdaq for at least 12 months. We purge the daily file of equity-days where there was no trading volume to eliminate the

Table 2 Descriptive Statistics on sample run-lengths and equity returns. This table presents cross-sectional statistics of run lengths and daily equity returns for quinquennial periods starting in January 1962 and ending in December 2005. A run is defined as an unreversed sequence of positive or negative returns. Cross-sectional statistics are calculated on the time-series mean and standard deviation of run lengths and equity returns over the quinquennial period. Cross-sectional statistics include the median and number of observations. The sample consists of all firms traded on the NYSE, Amex and Nasdaq. Mean run lengths tend to increase when there is a large increases in the number of listed stocks (as in the 1980s). When new listings are excluded from the data, by using stocks that span the entire length of the sample period, the run lengths drop back to levels slightly in excess of 2. For example, when firms with at least 30 years of data are used, the mean run lengths for the deciles are: $\{2.37, 2.39, 2.49, 2.47, 2.49, 2.41, 2.37, 2.20, 2.02\}$.

Period	Cross-Sectional Statistics	Run Length Mean	Run Length S Dev	Return (%) Mean	Return (%) S Dev
1962-1965	Mean	2.554	1.994	0.07462	1.65847
	Median	2.410	1.811	0.06406	1.50509
	N	3549	3549	3549	3549
1966-1970	Mean	2.479	1.908	0.04910	2.22603
	Median	2.404	1.804	0.03591	2.06242
	N	6650	6650	6650	6650
1971-1975	Mean	2.588	2.029	0.04425	2.27593
	Median	2.495	1.898	0.04381	2.16861
	N	4932	4932	4932	4932
1976-1980	Mean	2.531	1.953	0.10441	2.11741
	Median	2.471	1.863	0.08262	1.97729
	N	5868	5868	5868	5868
1981-1985	Mean	4.426	4.085	0.09026	2.19031
	Median	2.635	2.133	0.08847	2.05590
	N	10937	10937	10937	10937
1986-1990	Mean	3.948	3.440	0.04178	2.52132
	Median	2.505	1.972	0.04532	2.36671
	N	15190	15190	15190	15190
1991-1995	Mean	3.083	2.587	0.09645	2.77401
	Median	2.364	1.783	0.08514	2.53696
	N	18570	18570	18570	18570
1996-2000	Mean	2.509	2.029	0.09263	3.31840
	Median	2.230	1.641	0.07543	2.96127
	N	22185	22185	22185	22185
2001-2005	Mean	2.190	1.640	0.10244	2.75938
	Median	2.016	1.419	0.08293	2.40411
	N	17782	17782	17782	17782

possibility that a run is caused by stale data.[8] We then proceed to identify for each stock the length of every run-up and run-down which we define as a period of uninterrupted rise (drop) in stock price by using the equity returns including all distributions calculated in CRSP. Consistent with our model set up, days where there are zero returns for a given stock are assumed not to interrupt the current run (see the appendix which contains the algorithm). Run lengths are averaged for each security to create an average run length per stock for the period of interest.

[8] The absence of trading volume makes the comparison with other measures of illiquidity that are based on trading volume problematic. For example, the Amihud (2002) measure, where trading volume appears in the denominator of the illiquidity metric. We are also being conservative because this biases the results against run length being a proxy for illiquidity.

Table 2 reports descriptive statistics on the cross-section of the time-series mean and standard deviation of run lengths and equity returns for all stocks from January 1962 to December 2005, a period of forty-three years broken down into sub-periods of five years each, after an initial three-year sub-period.

Based on daily data, the mean run length varies from a low of 2.190 in the period 2001-05 to a high of 4.426 in the period 1981-85. Since the random walk hypothesis implies a mean run length of 2, there is some evidence that daily price paths are more persistent than random walks. In all periods median run lengths take values between 2 and 3. The ten year period from 1981 to 1990 appears to be different from the other years in the sample, since the run lengths are much higher on average, though not that much higher when we look at the medians. Therefore, the distribution of run lengths in the 1981-90 epoch is highly skewed, on account of the spurt in stocks listed on the NASDAQ.[9] We show that run length is a good proxy for liquidity, and hence, this period is characterized by marked liquidity imbalances in the cross-section of stocks. Heston and Sadka (2005) discover seasonality in liquidity, and here, we provide some evidence for liquidity regimes based on new stock listings.

4 Run Length and Trading Volume

We examine the relationship of our run length metric to trading volume in the cross-section of stocks. Note that the run length measure is computed from the price series of the stock and does not contain any trading volume information in its construction.

Figure 2 examines if there is a relationship between run length and trading volume using sorting into quintiles based on run length.[10] For each stock we computed the normalized daily trading volume (turnover) as the ratio of shares traded in a day to the outstanding shares in the firm. We are then able to calculate the trading volume (equally weighted across firms) in each quintile. We can see that the ordering by run length is similar to ordering by trading volume as per the analysis in Section 2. The relationship seen in Figure 2 is sharp and clear - as run length increases, trading volume declines (see the first set of bars on the left of the graph).

For completeness we also examined firm loadings on the Fama-French factors HML (High-minus-Low book-to-market), SMB (Small-minus-Big size) as well as the momentum factor UMD (Up-minus-Down). Figure 2 shows that sorting firms by these factor loadings does not provide a monotone ordering on trading volumes.[11]

Mindful of the argument that run length may simply be a proxy for autocorrelation in the returns, we computed the return autocorrelation for each stock and year. A cross-sectional regression (not tabulated) with Newey-West corrected standard errors of normalized trading volume on average run length *and* autocorrelation gives a negative

[9] When new listings are excluded from the data, by using stocks that span the entire length of the sample period, the run lengths drop back to levels slightly in excess of 2.

[10] Cognizant of the critique in Berk (2000), we keep the number of groups small, and only work with averages.

[11] There appears to be a *u*-shaped pattern, which suggests instead that trading volume is higher for stocks with extreme levels (high or low) of factor exposure. We may conjecture that the manner in which the factors are constructed might lead to this effect, i.e. they overweight the more frequently traded stocks because of their return characteristics. To the best of our knowledge, this link between trading volume and factor exposure has not been documented elsewhere. We do not know the underlying cause of this pattern, but it is tangential to the goals of this paper, and we leave it for further research.

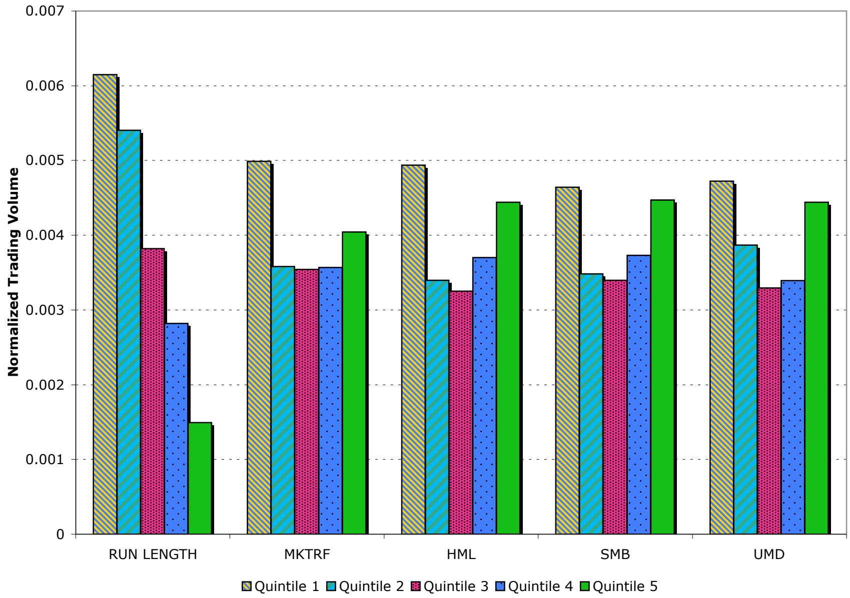

Fig. 2 Sorts of normalized trading volume by quintiles formed on run length and Fama-French factors. We sorted all firms into quintiles based on the desired variable, i.e. run length, beta coefficients on the standard asset-pricing factors: Excess market return, HML, SMB, and UMD. For each stock we computed the normalized daily trading volume as the ratio of shares traded in a day to the outstanding shares in the firm. We then calculate the trading volume (equally weighted across firms) in each quintile. The only monotone pattern for trading volume is in the run length quintiles. We see that trading volume declines from the lowest run length quintile to the highest.

relation to run length (t-statistic of 14.25) and a positive relation to auto-correlation (t-statistic of 22.74). If run length is left out of the regression the normalized trading volume is positively related to auto-correlation (t-statistic of 40.63). Thus, the inclusion of autocorrelation does not exclude run length as an explanatory variable for trading volume.

It is apparent from these analyses that sorting firms by run length is consistent with sorting firms by one aspect of liquidity, namely trading volume (turnover). In Figure 3, a more detailed depiction of the relationship between turnover and run length is shown. Consistent with our model, turnover is lower when run length increases.

5 Amihud Illiquidity

We have seen evidence that run lengths are related to trading volume in precisely the direction suggested by a liquidity relationship. In this section, we examine how run lengths relate to other measures of liquidity. A recent paper by Amihud (2002) develops a price impact measure of *illiquidity* for individuals stocks:

$$ILLIQ_{it} = \frac{1}{DAYS_{it}} \sum_{t=1}^{DAYS_{it}} \frac{|r_{it}|}{PRC_{it} \times VOL_{it}} \times 10^6,$$

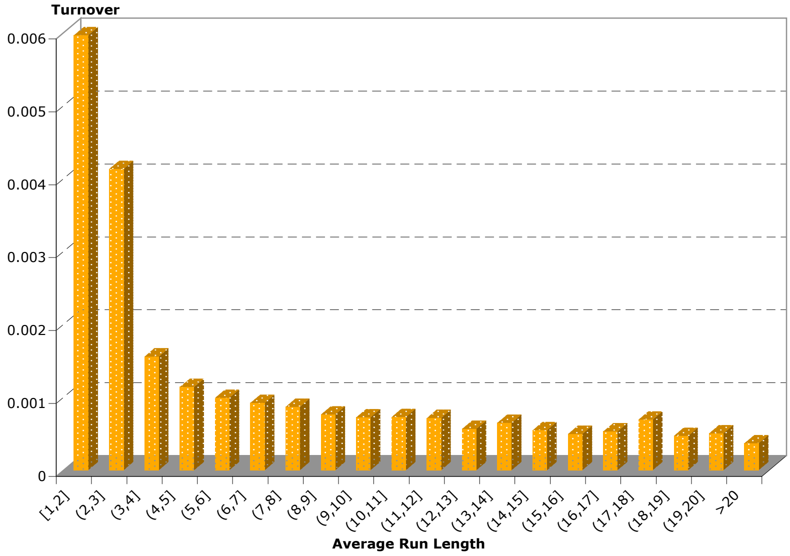

Fig. 3 Run length and turnover. The plot shows the relationship between normalized trading volume (turnover) as run length increases. The declining turnover evidences the fact that run length may be used as a proxy for illiquidity.

where r_{it} is the ith stock's return for day t, PRC_{it} is closing price, and VOL_{it} is daily trading volume, that is, the number of shares traded for a firm. $DAYS_{it}$ is the number of trading days for stock i in year t.[12] This proxy for liquidity is the same as used in the papers by Acharya and Pedersen (2005) who develop a liquidity-extended CAPM and Avramov, Chordia and Goyal (2006), who examine the relationship of liquidity to autocorrelation in stock returns. For a comparison of the spectral densities of returns and trading volumes, see Bonanno, Lillo and Mantegna (2000); the distribution of price impact is known to be power-law, per Lillo and Farmer (2005). For an alternative measure of price impact similar to the ideas here, see Pastor and Stambaugh (2003). After sorting firms into quintiles, we computed the average $ILLIQ$ measure within each quintile. Results are reported in Table 3.

We note that there is a relationship between the numerator and denominator of the $ILLIQ$ metric. Easley, Kiefer, O'Hara and Paperman (1996) have shown that the price impact of trades tends to be lower for frequently traded stocks, and thus this should result in a wider spread of $ILLIQ$ values; the earlier findings in Hasbrouck (1991) also suggest this result. This biases the results in Table 3 in favor of finding a relationship between run lengths and liquidity, given that we have already found a negative relationship between run lengths and trading volumes (in Figure 3). From the table we see that as run length increases, the measure of illiquidity also increases, supporting run length as a good proxy for liquidity. Further, the numbers in the table show that illiquidity increases rapidly in run length.

Amihud and Mendelson (1986) have shown that returns are correlated to bid-ask spreads, which are a measure of illiquidity. In our setting, we should find that run

[12] $ILLIQ$ is inversely proportional to the well-known "Amivest" liquidity measure.

Table 3 Run lengths and the Amihud measure. This table presents average Amihud illiquidity for groups determined by sorting on run length. This measure is as follows for each stock i: $ILLIQ_{it} = \frac{1}{DAYS_{it}} \sum_{t=1}^{DAYS_{it}} \frac{|r_{it}|}{PRC_{it} \times VOL_{it}} \times 10^6$, where r_{it} is the ith stock's return for day t, PRC_{it} is closing price, and VOL_{it} is trading volume, which is the number of shares traded for a firm. $DAYS_{it}$ is the number of trading days for stock i in year t. Normalized trading volume (turnover) is defined as trading volume divided by shares outstanding. The quoted bid-ask spread is defined as the difference between the ask and the bid prices divided by the average of the two. All measures are equally-weighted.

Run Quintile:	Lowest	2	3	4	Highest
ILLIQ:	0.2634	0.3470	0.4512	0.7144	3.1385
Turnover:	0.0062	0.0055	0.0039	0.0028	0.0015
Bid-Ask Spread:	0.0154	0.0209	0.0251	0.0286	0.0550

lengths are positively correlated to bid-ask spreads. Hence, we report bid-ask spreads for the run length quintiles in Table 3 as well. As expected, the results support the Amihud-Mendelson hypothesis.

We consider whether the relation between our illiquidity measure and Amihud's remains if we control for stock returns. We do this to make sure that differences in realized growth rates in the cross-section of stocks are not driving the differences in run lengths by injecting skewness in returns. We conduct a two-way 5×5 sort in which we first sort stocks by return into quintiles, and then within quintiles, sort stocks by run length. In each of the 25 cells in Table 4, we report the average of Amihud's illiquidity measure. We can see that the sort by our run length illiquidity measure lines up in the same way as does Amihud's measure, even after pre-sorting by contemporaneous returns (Panel A) or by the standard deviation of returns (Panel B).

We next explore whether run length remains a good proxy for illiquidity after including other possible controls. To assess this, we regressed illiquidity measures on average run length, and controlled for autocorrelation, trading volume and the interaction between the two, the results of which are presented in Table 5. Roll (1984) shows that under market efficiency, the bid-ask bounce induces negative autocorrelation (the sign is indeed negative in our regression). After controlling for this, we find that bid-ask spreads are positively correlated to run lengths, which are proxies for illiquidity (which enters with a positive coefficient).

Chordia, Roll and Subrahmanyam (2005) find that the capacity of an asset market to accommodate order imbalances, i.e. provide liquidity, is inversely related to the predictability of returns from previous order flows. Greater predictability, intuitively proxied by longer runs, is consistent with increases in illiquidity. Avramov, Chordia and Goyal (2006) find that high turnover (high NORMVOL) stocks in which there are short-run reversals (low AUTOCORR) are likely to be illiquid (measured as bid-ask spread), even after controlling for trading volume. Hence, we included an interaction term between trading volume and autocorrelation in the following regression. Each variable was demeaned by the firm's average value in order to account for fixed-effects. Mean run length remains a significant explanatory variable for illiquidity, whether measured by bid-ask spreads or the Amihud illiquidity metric.

In Table 5 we also control for the Lesmond, Ogden and Trzcinka (1999) zero return days measure since by construction it is closely associated to mean run lengths because

Table 4 The relation of Amihud's illiquidity measure with our run length measure. We examine illiquidity after controlling for contemporaneous stock returns (Panel A) and the contemporaneous standard deviation of stock returns (Panel B). The table reports a two-way 5 × 5 sort in which we first sort stocks by return (standard deviation of returns) into quintiles, and then within quintiles, sort stocks by run length. In each of the 25 cells in Table 4, we report the average of Amihud's illiquidity measure. We can see that the sort by our run length illiquidity measure lines up in the same way as does Amihud's measure, even after pre-sorting by returns or by standard deviation of returns.

Panel A: Stock Returns					
	Run Quintile				
	Lowest	2	3	4	Highest
Average Return Quintile					
1	0.280	0.333	0.430	0.652	3.187
2	0.234	0.302	0.378	0.697	3.373
3	0.203	0.312	0.380	0.657	2.918
4	0.238	0.332	0.436	0.647	2.956
5	0.389	0.469	0.637	0.992	3.353
Panel B: Standard Deviation of Stock Returns					
	Run Quintile				
	Lowest	2	3	4	Highest
StdDev of Return Quintile					
1	0.178	0.251	0.312	0.793	3.651
2	0.212	0.253	0.355	0.576	3.040
3	0.276	0.323	0.433	0.614	2.896
4	0.287	0.410	0.488	0.732	2.723
5	0.397	0.474	0.653	1.034	3.484

Table 5 Explaining liquidity with run length. The table presents results of regressions of two measures of liquidity, bid-ask spreads (BIDASK) and the Amihud illiquidity (ILLIQ) measure. After controlling for trading volume (NORMVOL), auto-correlation (AUTOCORR) in returns, an interaction term (NORMVOL*AUTOCORR), and separately we include the number of zero-return days (ZERORET) as in Lesmond, Ogden and Trzcinka (1999). In all cases the role of run length (MEANRUNLEN) as a proxy for illiquidity remains strongly statistically significant. In order to account for firm fixed-effects we demean each variable by the firm average. T-statistics estimated using GMM with the Newey and West (1987) correction with 2 lags are reported below the parameter estimates.

	Dependent variable			
	BIDASK		ILLIQ	
MEANRUNLEN	0.2667	0.6271	0.0014	0.0086
	8.21	25.54	5.74	43.06
AUTOCORR	-3.2877	-4.53892	-0.03432	-0.05993
	-24.35	-37.39	-36.15	-58.8
NORMVOL	-20.8808	-40.5924	-0.2116	-0.50297
	-3.85	-7.14	-8.78	-10.38
NORMVOL*AUTOCORR	246.3357	300.1359	1.4296	2.4606
	8.76	10.17	10.54	10.71
ZERORET	0.0166		0.0003	
	25.22		60.73	
R^2	11.99%	9.80%	36.76%	26.47%
N	66,732		105,200	

we let a run continue its course with zero returns (both metrics are complementary measures of illiquidity). Mean run length remains positive and significant.

6 Liquidity and Asset Pricing

6.1 Liquidity Factor Mimicking Portfolio

In this section we examine whether mean run length, as a proxy for illiquidity, is able to explain stock returns. First, we construct five portfolios at the end of each December from 1963 to 2005 based on quintiles of the average run-length. Eligible stocks are common stocks with a year-end stock price between \$5 and \$1000, trading on the NYSE, AMEX and Nasdaq for at least 12 months. Equally-weighted portfolio returns are then calculated for every month of the following year before rebalancing occurs[13]. We focus on equal-weighted run length portfolios as opposed to value-weighted because the range of returns across run length quintiles is smaller.[14]

We then construct a liquidity factor (denoted RLI) based on the run length portfolios. This factor is a mimicking portfolio consisting of the difference in returns between the quintile of longest run stocks (least liquid) and the quintile of shortest run stocks (most liquid).

$$
\begin{aligned}
RLI = & \text{ Highest Run-Length Portfolio Return} \\
& -\text{Lowest Run-Length Portfolio Return.}
\end{aligned}
$$

We plotted the time series of this liquidity measure in Figure 4. From the figure one can see that the factor shows a mean positive level, with frequent episodes where the factor spikes into positive or negative regions. There is a period in the mid-70s when the illiquidity factor became highly variable, as well as in the recent period during the crisis during and following the failure of the dot-com firms.

We examine whether our liquidity factor can explain the cross-section of asset returns. We formed 25 portfolios based on size and book-to-market quintiles (these

[13] We adjust returns for stock delistings to account for the possibility that the stocks in the most illiquid portfolio may be more likely to leave the sample in the post-formation months thereby causing a survivorship bias. In particular, Shumway (1997) finds that stocks with smaller market capitalization are more susceptible to the delisting bias and the delisting announcement tends to be preceded by a large spike in volume turnover. Since both effects are likely to disproportionally affect the returns on the most illiquid portfolio we adjust returns following Shumway (1997) for NYSE and Amex stocks and Shumway and Warther (1999) for those listed on Nasdaq. More precisely, the last return used is the last return available on CRSP, or the delisting return if available. If the stock delisting was coded as 500 (reasons unavailable), 520 (went to OTC), 551-573 and 580 (various reasons), 574 (bankruptcy) and 584 (does not meet exchange financial guidelines) and the delisting return is missing we assign a return of -30% for companies listed on the NYSE or AMEX and -55% for companies listed on Nasdaq.

[14] Several studies focus on equal-weighted return and illiquidity measures, for instance Amihud (2002) and Chordia, Roll and Subrahmanyam (2005). Computing market returns and illiquidity as equal-weighted averages is a way of compensating for the overrepresentation in our sample of large liquid securities, as compared to the true market portfolio in the economy. In particular, our sample does not include illiquid assets such as corporate bonds, private equity, real estate, and many small stocks, and these assets constitute a significant fraction of aggregate wealth. Therefore, we focus in our empirical work on an equal-weighted market portfolio, although we also estimate the model with a value-weighted market portfolio for robustness (as in Acharya and Pedersen (2005)).

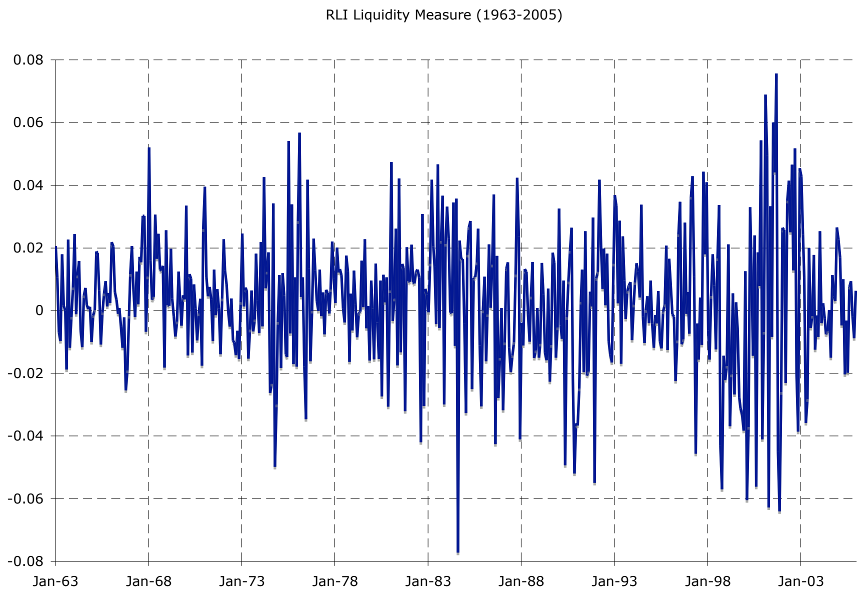

Fig. 4 Liquidity mimicking portfolio. Time series plot of the monthly cross-sectional average return of a portfolio constructed from positions in all stocks in the top quintile of run length minus the average return of a portfolio of stocks in the bottom run length quintile.

portfolios are obtained from Kenneth French's web site). For each portfolio we ran a time series regression of equal-weighted portfolio returns on the Fama-French factors and our run length measure of illiquidity (RLI). Using the coefficients in the time series regressions we then ran factor-pricing cross-sectional regressions. Time series coefficients were obtained using rolling five-year windows of monthly data for the first pass regression, followed by a second pass cross-sectional regression using the returns in the month immediately after the rolling window. Therefore, we rolled forward a month at a time resulting in 456 cross-sectional regressions. Table 6 contains the results, which show that our illiquidity measure is significant even in the presence of the other asset pricing factors. Thus, analogous to the finding in Acharya and Pedersen (2005) and Korajczyk and Sadka (2006), the CAPM model is extendable with a liquidity factor and we are able to assess the differential explanatory power for large versus small stocks. In Panels A and B of Table 6, we report the results of the cross-sectional regression when the size and book-to-market portfolios are equal-weighted and value-weighted respectively. As a check, we also redid these regressions by forming 25 liquidity portfolios based on run lengths and the results remain unchanged. These are reported in Panels C and D. Our liquidity mimicking portfolio (RLI) is significant in all cases when equal-weighted portfolios are used, and only mildly so for value-weighted portfolios. Value-weighting overemphasizes large stocks where liquidity is less likely to explain returns. Overall, RLI remains a significant explanatory variable of the cross-section of stock returns even after controlling for other asset pricing factors. Interestingly, the momentum factor (UMD) does not appear to explain the cross-section of asset prices when the RLI factor is included. These results are also robust to whether the equal-weighted or value-weighted market portfolio is used. Additionally, we split the sample

Table 6 Cross-sectional asset pricing regressions for the run length factor. In Panel A and B, 25 portfolios are formed on size and B/M following the methodology of Fama and French (1993). In Panel C and D, 25 portfolios are formed on average run length (liquidity portfolios). Each portfolio's excess returns are then regressed on the excess equal-weighted market returns (EWMKTRF), SMB, HML and UMD and the equally weighted liquidity mimicking portfolio (RLI). RLI is constructed as the returns on the portfolio based on the highest quintile of average run lengths minus the corresponding lowest quintile. The portfolios excess returns are then regressed every month against the beta coefficients estimated in the previous 60 months. The time-series of coefficients are then averaged and T-statistics are calculated using GMM with the Newey and West (1987) correction with 2 lags. Three regressions are run and reported in each of the four panels below: *(a)* Pure CAPM extended by RLI, *(b)* Fama-French extended, and *(c)* Carhart extended. In addition, the coefficients for each model for two sub-periods January-1969 to June-1987 (RLI1)and July-1987 to December-2005 (RLI2)are presented only for the RLI measure . T-statistics are reported below the coefficients.

	Panel A: Equally-weighted Size and BM portfolios							
	α	$\beta_{EWMKTRF}$	β_{HML}	β_{SMB}	β_{UMD}	β_{RLI}	β_{RLI1}	β_{RLI2}
(a)	0.01258	-0.00535				0.00442	0.004464	0.004368
	3.98	-1.28				3.27	2.40	2.22
(b)	0.01518	-0.00822	0.00441	0.00089		0.00351	0.00244	0.004625
	6.20	-2.54	2.49	0.54		2.81	1.31	2.77
(c)	0.01471	-0.00778	0.00423	0.00135	0.00338	0.00355	0.002857	0.004276
	6.01	-2.43	2.40	0.82	1.05	3.06	1.61	2.86
	Panel B: Value-weighted Size and BM portfolios							
	α	$\beta_{EWMKTRF}$	β_{HML}	β_{SMB}	β_{UMD}	β_{RLI}	β_{RLI1}	β_{RLI2}
(a)	0.01029	-0.00332				0.00332	0.003337	0.003297
	3.43	-0.81				2.22	1.69	1.47
(b)	0.0142	-0.0083	0.0044	0.0011		0.0016	0.000261	0.002982
	5.36	-2.43	2.61	0.70		1.43	0.17	1.86
(c)	0.01350	-0.00839	0.00447	0.00116	0.00196	0.00161	0.00037	0.002876
	4.77	-2.39	2.70	0.72	0.68	1.45	0.23	1.84
	Panel C: Equally-weighted liquidity portfolios							
	α	$\beta_{EWMKTRF}$	β_{HML}	β_{SMB}	β_{UMD}	β_{RLI}	β_{RLI1}	β_{RLI2}
(a)	0.00920	-0.00309				0.00297	0.003552	0.002392
	3.06	-0.86				2.62	2.35	1.41
(b)	0.00935	-0.00327	0.00365	-0.00314		0.00304	0.003485	0.002615
	2.73	-0.89	1.68	-1.33		2.81	2.53	1.57
(c)	0.00936	-0.00363	0.00463	-0.00403	0.00303	0.00277	0.00333	0.002224
	2.65	-0.95	2.13	-1.86	0.89	2.53	2.42	1.30
	Panel D: Value-weighted liquidity portfolios							
	α	$\beta_{EWMKTRF}$	β_{HML}	β_{SMB}	β_{UMD}	β_{RLI}	β_{RLI1}	β_{RLI2}
(a)	0.00672	0.01298				0.00267	0.003167	0.002164
	2.24	2.98				2.02	1.66	1.18
(b)	0.00969	0.00826	0.00018	0.00738		0.00247	0.00341	0.001537
	3.30	2.15	0.10	3.54		1.85	1.92	0.77
(c)	0.00903	0.00894	0.00064	0.00788	0.00278	0.00218	0.002884	0.001428
	3.00	2.24	0.34	3.61	0.83	1.57	1.61	0.67

into two sub-periods with 222 observations each (January 1969-June 1987 and July 1987-December 2005) and estimated the two models separately. We only report the coefficient and significance on the RLI variable. The sign of the RLI coefficient remains ualtered, though the significance of the coefficients is reduced in some subperiods since there are now only 222 months in the regression. The coefficients for the entire period and the two subperiods are of similar magnitude. Thus, the effect of run length is consistent through the subperiods.

Since previously we found that the number of zero return days in a year of a stock appear to be a complementary measure of illiquidity, we examine whether a mimicking portfolio based on that measure is able to explain the cross-section of asset prices. We create the mimicking portfolio based on zero return days (ZERORET) similarly to the RLI factor. We then include the ZERORET factor with the RLI factor in the Fama and MacBeth (1973) regressions. The results are reported in Table 7. As expected, ZERORET is important in explaining the cross-section of stock returns and it is complementary to the RLI factor since this latter retains its significance. The ZERORET factor, however, appears to be more robust to the choice of test portfolios (size/book-to-market or liquidity portfolios) and their weighting scheme (equal- or value-weighted). We believe that this is the first demonstration that the number of zero return trading days, earlier used for representing transaction costs in Lesmond, Ogden and Trzcinka (1999), is a priced illiquidity factor.[15] Again, as a robustness check, we split the sample into two sub-periods with 222 observations each (January 1969-June 1987 and July 1987-December 2005) and estimated the two models separately. Though the significance on the coefficients is reduced in some subperiods since there are now only 222 months in the regressions, the coefficients for the entire period and the two subperiods are of similar magnitude. The effect of run length is consistent through the subperiods.

6.2 Characteristic Regressions

In this section, we test the asset pricing implications of the average run length by using it as a characteristic in cross-sectional regressions of individual stock excess returns following Amihud and Mendelson (1986). We compute excess returns for each security-month using the CRSP value-weighted index.[16] For each stock-month of year t we calculate the following characteristics:

1. The average run length computed from January to December of year $(t-1)$ (denoted MEANRUNLEN).
2. The natural logarithm of the ratio of the book value of equity plus deferred taxes to the market value of equity (BM/ME), using the end of the previous year market and book values. As in Fama and French (1992), the value of BM for July of year t to June of year $(t+1)$ was computed using COMPUSTAT data at the end of year $(t-1)$.
3. The logarithm of the market value of the equity of the firm as the end of the previous year (SIZE).
4. The cumulative return over the two months ending at the beginning of the previous month (RET(2-3)).
5. The cumulative return over the three months ending three months previously (RET(4-6)).
6. The 6 month cumulative return ending 6 months previously (RET(7-12)).

Every month we estimate cross-sectional regressions of individual stock excess returns on the characteristics. The monthly time-series of coefficients is then averaged

[15] We expect that the zero volume measure of Liu (2003) will provide similar results. When we computed the correlation between the monthly firm-average of the zero volume measure with ZERORET, we obtained a correlation of 0.88 with a p-value less than 0.0001.

[16] Similar results are obtained if the CRSP equally-weighted index is used instead.

Table 7 Cross-sectional asset pricing regressions for the components of the run length factor. In Panel A and B, 25 portfolios are formed on size and B/M following the methodology of Fama and French (1993). In Panel C and D, 25 portfolios are formed on average run length (liquidity portfolios). Each portfolio's excess returns are then regressed on the excess equal-weighted market returns (EWMKTRF), SMB, HML and UMD and the illiquidity mimicking portfolios (ZERORET and RLI). ZERORET (RLI) is constructed as the returns on the portfolio based on the highest quintile of the number of zero return trading days (mean run lengths) in a year minus the corresponding lowest quintile. The portfolios excess returns are then regressed every month against the beta coefficients estimated in the previous 60 months. The time-series of coefficients are then averaged and T-statistics are calculated using GMM with the Newey and West (1987) correction with 2 lags. Three regressions are run and reported in each of the four panels below: *(a)* Pure CAPM extended by ZERORET and RLI, *(b)* Fama-French extended, and *(c)* Carhart extended. In addition, the coefficients for each model for two sub-periods January-1969 to June-1987 (RLI1) and July-1987 to December-2005 (RLI2) are presented only for the RLI measure. T-statistics are reported below the coefficients.

	α	$\beta_{EWMKTRF}$	β_{HML}	β_{SMB}	β_{UMD}	$\beta_{ZERORET}$	β_{RLI}	β_{RLI1}	β_{RLI2}
Panel A: Equally-weighted Size and BM portfolios									
(a)	0.011241	-0.004086				0.008036	0.003764	0.003698	0.00333
	4.59	-1.17				3.62	2.61	2.06	1.69
(b)	0.016143	-0.009187	0.004402	0.000976		0.007418	0.004905	0.003699	0.006133
	6.59	-2.73	2.55	0.60		3.63	3.80	1.93	3.56
(c)	0.015275	-0.008186	0.004255	0.001583	0.004141	0.007224	0.005103	0.003343	0.006894
	6.24	-2.58	2.49	0.97	1.30	3.60	3.83	1.87	3.52
Panel B: Value-weighted Size and BM portfolios									
	α	$\beta_{EWMKTRF}$	β_{HML}	β_{SMB}	β_{UMD}	$\beta_{ZERORET}$	β_{RLI}	β_{RLI1}	β_{RLI2}
(a)	0.008911	-0.002480				0.006909	0.001972	0.001636	0.002307
	3.71	-0.69				2.89	1.41	1.04	0.99
(b)	0.014828	-0.008966	0.004489	0.001212		0.004558	0.002343	0.001223	0.003469
	5.60	-2.52	2.70	0.76		2.26	1.86	0.73	1.85
(c)	0.013596	-0.008580	0.004482	0.001139	0.000797	0.004350	0.002175	0.001089	0.003269
	4.81	-2.45	2.72	0.72	0.29	2.18	1.75	0.66	1.76
Panel C: Equally-weighted liquidity portfolios									
	α	$\beta_{EWMKTRF}$	β_{HML}	β_{SMB}	β_{UMD}	$\beta_{ZERORET}$	β_{RLI}	β_{RLI1}	β_{RLI2}
(a)	0.010129	-0.004042				0.005007	0.002629	0.003	0.002244
	3.20	-1.18				2.66	2.31	2.18	1.23
(b)	0.009775	-0.003410	0.003677	-0.004931		0.004952	0.001828	0.00253	0.001063
	2.82	-0.94	1.66	-2.22		2.62	1.62	1.93	0.58
(c)	0.009857	-0.003349	0.004050	-0.004902	0.002210	0.004983	0.001439	0.002525	0.000277
	2.79	-0.87	1.87	-2.41	0.54	2.62	1.24	1.87	0.15
Panel D: Value-weighted liquidity portfolios									
	α	$\beta_{EWMKTRF}$	β_{HML}	β_{SMB}	β_{UMD}	$\beta_{ZERORET}$	β_{RLI}	β_{RLI1}	β_{RLI2}
(a)	0.008854	0.010083				0.004581	0.002016	0.00351	0.000523
	3.22	2.69				2.19	1.38	1.99	0.23
(b)	0.009669	0.008692	0.000447	0.007517		0.003695	0.002434	0.00332	0.001551
	3.35	2.33	0.22	3.66		1.63	1.65	1.88	0.66
(c)	0.008985	0.009194	0.001086	0.007624	0.003862	0.003912	0.002112	0.002562	0.001635
	3.01	2.28	0.53	3.55	1.13	1.66	1.40	1.43	0.67

as in Fama and MacBeth (1973). T-stats are obtained using GMM with the Newey and West (1987) correction with 2 lags. The BM/ME and SIZE characteristics control for the book-to-market and size effects, respectively, while RET(2-3), RET(4-6) and RET(7-12) control for the momentum effect. The results are reported in table 8. As predicted by theory, the MEANRUNLEN coefficients are positive in all model specifications. When run length is the only characteristic (model 1), the coefficient is positive and highly significant. Furthermore, run length continues to be highly significant when

Table 8 Cross-sectional characteristic regressions. Coefficients estimates are the time-series averages of monthly individual stock cross-sectional OLS regressions. The dependent variable is the individual stock excess returns calculated using the CRSP value-weighted index. BM/ME is the natural logarithm of the ratio of the book value of equity plus deferred taxes to the market value of equity, using the end of the previous year market and book values. As in Fama and French (1992), the value of BM for July of year t to June of year $t+1$ was computed using COMPUSTAT data at the end of year $t-1$. SIZE is the logarithm of the market value of the equity of the firm as the end of the previous year. RET(2,3) is the cumulative return over the two months ending at the beginning of the previous month. RET(4,6) is the cumulative return over the three months ending three months previously. RET(7-12) is the 6 month cumulative return ending 6 months previously. In addition, the coefficients for each model for two sub-periods January-1963 to December-1985 and January-1986 to December-2005 are presented only for the RLI measure. T-statistics are calculated using GMM with the Newey and West (1987) correction with 2 lags.

	Model 1		Model 2		Model 3		Model 4	
INTERCEPT	-0.00272	*	-0.00071		0.00574		0.00321	
	-1.67		-0.42		1.54		0.98	
MEANRUNLEN	0.00232	***	0.00195	***	0.00097	*	0.00090	*
	4.00		3.41		1.74		1.76	
BM/ME			0.00302	***	0.00210	**	0.00243	***
			3.72		2.36		3.03	
SIZE					-0.00083	*	-0.00077	*
					-1.88		-1.93	
RET(2-3)							0.00678	**
							2.12	
RET(4-6)							0.00901	***
							3.22	
RET(7-12)							0.01014	***
							5.00	
MEANRUNLEN coefficients for sub-period regressions								
196301-198412	0.004635	***	0.004152	***	0.002595	***	0.002278	***
	4.93		4.33		3.11		2.91	
198501-200512	0.00000		-0.00026		-0.00065		-0.00048	
	0.01		-0.50		-0.93		-0.78	

Significance: 1%: ***, 5%: **, 10%: *

the BM/ME characteristic is added (model 2). When SIZE is also included (model 3), run length is significant at the 10% level, and it remains significant even after momentum effects are added (model 4). These results complement those of the previous section and confirm the role of run length as a priced characteristic and underline its role in asset pricing.

6.3 Connection to other Liquidity Measures

We explore the commonality across different liquidity measures using a principal component analysis (PCA) in a manner similar to that in Korajczyk and Sadka (2006). Here, we take sample of 6 liquidity measures, including our own, and determine how similar these measures are. Our goal here is to see whether our measure resides within the space of the other measures or appears to be somewhat different in its behavior.

We use the following 6 monthly time-series of liquidity: turnover, bid-ask spread, the Amihud illiquidity measure, and the number of zero return days, all averaged across securities within each month. In addition, we include the Pastor and Stambaugh (2003) liquidity innovations. Our own measure of mean run length is averaged across firms

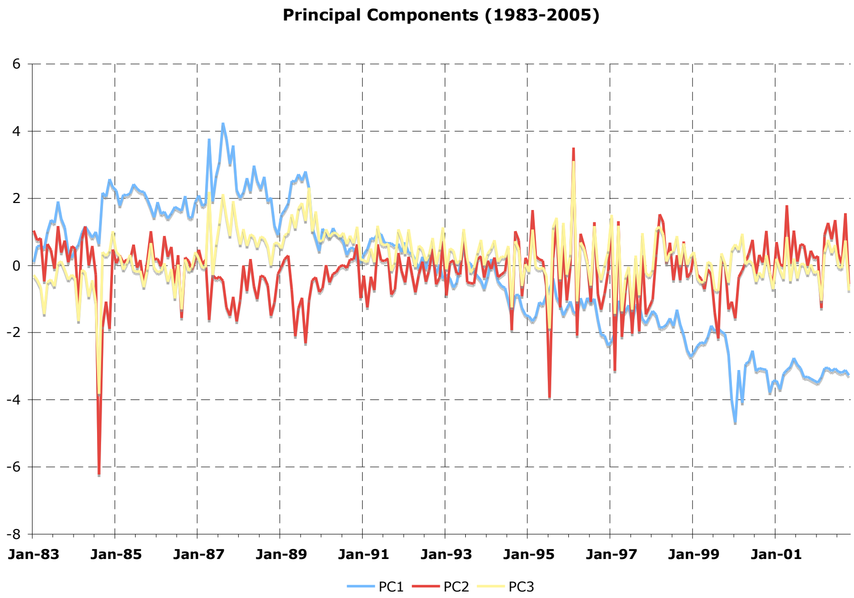

Fig. 5 Time series plot of the first three principal components of the liquidity measures. See Table 9.

each month. These six liquidity measures are then subjected to principal components analysis.

Table 9 presents the results of the analysis. We show the correlation matrix of the factors and the results of the PCA in Panel A. Our measure of mean run length is negatively correlated to turnover as is to be expected. The measure is positively correlated to bid-ask spreads, the zero-return illiquidity measure and Amihud's illiquidity metric. While all these correlations are statistically significant, the correlation with Pastor and Stambaugh's innovation measure is not.

Panel B of Table 9 shows that the first principal component explains 61% of the common variation amongst the measures. The second and third components explain 17% and 15% of the variation (Figure 5 shows the time series of the first thee components). Therefore, there appears to be a well-defined space spanned by the various measures of liquidity. Panel C of the table shows the eigenvectors of the components, detailing the loadings of each liquidity measure on the components (note that the sign of the loading need not be consequential). Except for the PS measure, the loadings of all the other five liquidity measures are similar for the first principal component. The magnitude and pattern of the loadings does suggest that our mean run length measure of liquidity demonstrates behavior similar to that of the other measures, and offers an excellent complementary measure especially when the computation of the other measures is more onerous. The fact that the run length measure only requires the times series of stock prices, and not any other data such as trading volume, makes it an easier metric to implement for practical purposes. Unlike some measures that need regressions analysis for their construction, the run length measure requires no pre-processing or

Table 9 Principal components analysis of liquidity measures. We explore the commonality across different liquidity measures using a principal component analysis (PCA). In the PCA we use 6 monthly time-series of liquidity: bid-ask spread (BASPREAD), turnover (TURN), the Amihud illiquidity measure (ILLIQ), and the number of zero return days (ZERORET) are averaged monthly across securities. We also include the Pastor and Stambaugh (2003) liquidity innovations (PSINNOV). The measure of mean run length (MEANRUNLEN) is averaged across firms each month. The three panels below present the correlation matrix of the liquidity measures, the eigenvalues of the correlation matrix, and the loadings or eigenvectors. In the correlation matrix, each cell contains two numbers. The number on top is the correlation, and the number below it is the p-value. The number of paired observations in each cell used to compute the correlation is the number of months of data, i.e. 276.

Panel A: Correlation Matrix

	MEANRUNLEN	TURN	BASPREAD	ZERORET	ILLIQ
TURN	-0.7909				
	.0001				
BASPREAD	0.3810	-0.5896			
	.0001	.0001			
ZERORET	0.8736	-0.8729	0.6786		
	.0001	.0001	.0001		
ILLIQ	0.3909	-0.5477	0.8444	0.6074	
	.0001	.0001	.0001	.0001	
PSINNOV	0.0603	-0.0789	-0.0150	0.0883	0.0153
	0.3183	0.1913	0.8042	0.1437	0.8005

Panel B: Eigenvalues of the Correlation Matrix

	Eigenvalue	Proportion	Cumulative
1	3.6549	0.6091	0.6091
2	1.0461	0.1744	0.7835
3	0.8961	0.1494	0.9329
4	0.1967	0.0328	0.9656
5	0.1575	0.0262	0.9919
6	0.0487	0.0081	1.0000

Panel C: Eigenvectors of the Principal Components

	PC1	PC2	PC3
MEANRUNLEN	0.4264	0.2688	-0.4634
TURN	-0.4707	-0.1427	0.2167
BASPREAD	0.4231	-0.3386	0.4155
ZERORET	0.4984	0.1216	-0.1758
ILLIQ	0.4094	-0.3226	0.4607
PSINNOV	0.0401	0.8209	0.5679

statistical analysis. It is very simple to compute (see the parsimony of the algorithm in Appendix A).

7 Conclusions

Based on a simple model of stock price changes, we derive that a stock's average run length, irrespective of observation frequency, is inversely proportional to trading volume and directly proportional to trade price impact. Hence, we present a theoretical model for why increasing run length corresponds to increasing illiquidity. We showed, over a period of forty-five years, that sorting stocks by run length is consistent with sorting by standard liquidity measures, such as turnover, bid-ask spreads, and other price-impact measures such as that of Amihud (2002). Using standard asset-pricing tests, we find that our constructed liquidity factor is priced in the cross-section of asset returns. It is also shown to be a priced firm characteristic. As an added bonus, our price-based measure enables us to construct proxies for liquidity even when data on trading volume is unavailable, as is often the case with international data. We also find that the zero-return days measure of Lesmond, Ogden and Trzcinka (1999) is a priced factor, and we believe this is the first analysis of this effect. The run length measure is not eliminated in regressions with the zero-return measure. A principal components analysis of liquidity metrics evidences a common space spanned by the measures. A comparison of these (and various other) illiquidity measures on various performance dimensions is an interesting avenue for further research.

Our theoretical framework is a simple one that naturally encompasses liquidity measures and connects price patterns to liquidity. We view this as explaining the underpinning statistical mechanics of illiquidity. This research may be furthered in many ways. First, there is a growing literature that documents seasonal effects in liquidity [see Heston and Sadka (2005) for one example], and we can examine whether these effects are evident in the run length measure. Second, illiquidity is persistent, and further investigations of the sources of this persistence are predicated, especially if they help explain persistence in returns through the momentum effect [Jegadeesh and Titman (1993)]. Third, another strand of the literature (see Raghubir and Das (2004)) finds that investors evidence preferences for liquidity by preferring stocks of shorter average run length, hence subconsciously showing liquidity preference. Fourth, there is growing evidence that equity market liquidity factors may be useful also in explaining credit default swap spreads (see de Jong and Driessen (2005), Das and Hanouna (2007)), and hence the run length measure that is easy to compute would be immediately applicable in those markets. And finally, trading strategies based on run length may be used to assess if liquidity premia are correctly priced.

A Appendix: Run Length Algorithm

A run comprises a string of returns without a reversal. It is possible that we obtain a string of up (down) moves followed by zero returns, and then more up (down) moves. The presence of zeros does not terminate the run. If the zeros are followed by a different sign return from the one preceding the zeros, then the run terminates with the last zero. The shortest run length is 1. The pseudo-code for the algorithm is as follows:

1. Denote the string of returns as $\{x_1, x_2, \ldots, x_n\}$.
2. Initialize: `FreqTable of Run Lengths`, i.e. $f(RunLen)$, $s = sign(x_1)$, $RunLen = 1$.
3. For i in (2,n):
    - If $sign(x_i) \times s \geq 0$, $RunLen++$
    - If $sign(x_i) \times s < 0$
        - $f(RunLen)++$
        - $RunLen = 1$
        - $s = sign(x_i)$

References

Acharya, Viral., and Lasse Pedersen (2005). Asset Pricing with Liquidity Risk, *Journal of Financial Economics* 77(2), 375-410.

Amihud, Yakov. (2002). Illiquidity and Stock Returns: Cross-Section and Time-Series Effects, *Journal of Financial Markets* 5, 31-56.

Amihud, Yakov., and Haim Mendelson (1986). Asset Pricing and the Bid-Ask Spread, *Journal of Financial Economics* 17, 223-249.

Amihud, Yakov., Haim Mendelson, and Lasse Pedersen (2005). Liquidity and Asset Prices, *Foundations and Trends in Asset Prices* 1(4), 269-364.

Avramov, Doron., and Tarun Chordia (2001). Characteristic Scaled Betas, Working paper, Emory University.

Avramov, Doron., Tarun Chordia, and Amit Goyal (2006). Liquidity and Autocorrelations in Individual Stock Returns, *Journal of Finance* 61(5), 2365-2394.

Berk, Jonathan (2000). Sorting out sorts, *Journal of Finance* 55, 407-427.

Bonanno, Giovanni., Fabrizio Lillo, and Rosario Mantegna (2000). Dynamics of the Number of Trades of Financial Securities, *Physica* A 280, 136-141.

Brennan, Michael., and Avanidhar Subrahmanyam (1996). Market Microstructure and Asset Pricing: On the Compensation for Illiquidity in Stock Returns, *Journal of Financial Economics* 41, 441-464.

Campbell, John., Sanford Grossman, and Jiang Wang (1993). Trading Volume and Serial Correlation in Stock Returns, *Quarterly Journal of Economics* 108, 905-939.

Carhart, Mark (1997). On Persistence in Mutual Fund Performance, *Journal of Finance* 52(1), 57-82.

Cetin, U., Robert Jarrow, Philip Protter, and M. Warachka (2006). Pricing Options in a Black-Scholes Economy with Illiquidity: Theory and Empirical Evidence, *Review of Financial Studies* 19(2), 493-529.

Chacko, George (2004). Liquidity Risk in the Corporate Bond Markets, Working paper, Harvard Business School.

Chacko, George., Jakub Jurek and Erik Stafford (2008). The Price of Immediacy, *Journal of Finance*.

Chacko, George., Sriketan Mahanti, Gaurav Mallik, and Marti Subrahmanyam (2005). The Determinants of Liquidity in the Corporate Bond Markets: An Application of Latent Liquidity, forthcoming, *Journal of Financial Economics*.

Chordia, Tarun., Richard Roll, and Avanidhar Subrahmanyam (2005). Liquidity and Market Efficiency, Working paper, Emory University.

Chordia, Tarun, Amit Goyal, Gil Sadka, Ronnia Sadka, and Lakshmanan Shivakumar (2006). Liquidity and the Post-Earnings-Announcement-Drift, Working paper, Emory University.

Conrad, Jennifer., Alaudeen Hameed, and Cathy Niden (1994). Volume and Autocovariances in Short-Horizon Individual Stock Returns, *Journal of Finance* 49(4), 1305-1329.

Daniel, Kent., and Sheridan Titman (1997). Evidence on the Characteristics of Cross-sectional Variation in Stock Returns, *Journal of Finance*,

Das, Sanjiv., and Paul Hanouna (2007). Hedging Credit: Equity Liquidity Matters, *Journal of Financial Intermediation*, forthcoming.

de Jong, Frank., and Joost Driessen (2005). Liquidity Risk Premia in Corporate Bond Markets, Working paper, University of Amsterdam.

Easley, David., Nicholas Kiefer, Maureen O'Hara, and J. Paperman (1996). Liquidity, Information, and Infrequently Traded Stocks, *Journal of Finance* 51, 1405-1436.

Easley, David., Nicholas Kiefer, and Maureen O'Hara (1997). One Day in the Life of a Very Common Stock, *The Review of Financial Studies* 10(3), 805835.

Edgington, Eugene, S. (1961). Probability table for number of runs of signs of first differences in ordered series, *Journal of the American Statistical Society* 56, 156-159.

Fama, Eugene. F., (1965). The Behavior of Stock Market Prices, *Journal of Business* 38(1), 34-105.

Fama, Eugene. F., and Kenneth R. French (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance* 47, 427-465.

Fama, Eugene. F., and Kenneth R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3-56.

Fama, Eugene. F., and James D. MacBeth (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 81, 607-636.

Getmansky, Mila., Andrew Lo, and Igor Makarov (2004). An Econometric Model of Serial Correlation and Illiquidity in Hedge Fund Returns, *Journal of Financial Economics* 74(3), 529-610.

Glosten, Lawrence, R. (1989). Insider trading, liquidity, and the role of the monopolist specialist, *Journal of Business* 62(2), 211-235.

Grafton, R.G.T (1981). The runs-up and runs-down tests, *Applied Statistics* 30(1), 81-85.

Grossman, Sanford., and Merton Miller (1988). Liquidity and Market Structure, *Journal of Finance* 43, 617-633.

Hasbrouck, Joel (1991). Measuring the Information Content of Stock Trades, *Journal of Finance* 46, 179-207.

Hasbrouck, Joel (2006). Trading costs and returns for US equities: Estimating effective costs from daily data, Working paper, New York University.

Hendershott, Terry., and Mark Seasholes (2007). Market Maker Inventories and Stock Prices, forthcoming *American Economic Review Papers and Proceedings*.

Heston, Steven., and Ronnie Sadka (2005). Seasonal Liquidity and Stock Returns, Working paper, University of Washington, Seattle.

Jegadeesh, Narasimhan., and Sheridan Titman (1993). Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jones, Charles., Gautum Kaul, and Marc Lipson (1994). Transactions, Volume and Volatility, *Review of Financial Studies* 7, 631-651.

Karpoff, Jonathan (1987). The Relation between Price Change and Trading Volume: A Survey, *Journal of Financial and Quantitative Analysis* 22, 109-126.

Korajczyk, Robert., and Ronnie Sadka (2006). Pricing the Commonality Across Alternative Measures of Liquidity, forthcoming *Journal of Financial Economics*.

Lesmond, David., Joseph Ogden, and Charles Trzcinka (1999). A New Estimate of Transaction Costs, *Review of Financial Studies* 12(5), 1113-1141.

Li, Jinliang., Robert Morradian, and Wei Zhang (2007). Is Illiquidity a Risk Factor? A Critical Look at Commission Costs, *Financial Analysts Journal* 63(4), 28-39.

Lillo, Fabrizio., and J. Doyne Farmer (2005). The Key Role of Liquidity Fluctuations in Determining Large Price Fluctuations, *Fluctuations and Noise Letters* 5, 209-216.

Liu, Weimin (2003). Liquidity Premium and A Two-factor Model, Working paper, Manchester Business School.

McQueen, Grant., and Steven Thorley (1994). Bubbles, Stock Returns and Duration Dependence, *Journal of Financial and Quantitative Analysis* 29(3), 379-401.

Moore, P. G., (1978), Nonparametric Statistics: Runs, in W. J. Kruskal and J. Tanur (eds.), *International Encyclopedia of Statistics*, Free Press, New York, 655-651.

Newey, W. and K. West (1987), A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55, 703-708.

Pagano, Marco (1989). Trading Volume and Asset Liquidity, *Quarterly Journal of Economics* 104, 255-274.

Pastor, Lubos. and Robert. F. Stambaugh (2003). Liquidity Risk and Expected Stock Returns. *Journal of Political Economy* 111, 642-685.

Raghubir, Priya., and Sanjiv Das (2004). The Long and Short of it: Why are stocks with shorter run-lengths preferred? Working paper, UC Berkeley.

Roll, Richard (1984). A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *The Journal of Finance* 39(4), 1127-1139.

Sadka, Ronnie (2006). Momentum and Post-Earnings- Announcement Drift Anomalies: The Role of Liquidity Risk, *Journal of Financial Economics* 80, 309-349.

Shumway, T. (1997), The Delisting Bias in CRSP Data. *Journal of Finance* 52, 327-340.

Shumway, T., and V.A. Warther (1999). The Delisting Bias in CRSPs Nasdaq Data and Its Implications for the Size Effect. *Journal of Finance*, 54, 2361-2379.

Wyss, Rico von (2004). Measuring and Predicting Liquidity in the Stock Market, Dissertation, University of St. Gallen.