

# Pricing Implications of Shared Variance in Liquidity Measures

Loran Chollete

Norwegian School of Economics and Business Administration, Norway

Randi Næs

Norges Bank, Norway

Johannes A. Skjeltorp

Norges Bank, Norway

March 7, 2006

## Abstract

This paper constructs fundamental liquidity measures and investigates the pricing implications of shared variation in a large set of high frequency liquidity measures. Through a common factor analysis we estimate three orthogonal, market-wide liquidity variables that statistically capture time series variations in market wide liquidity. We also test for differential explanatory power of trade-based and order-based factors. We uncover two main results. First, after estimating securities' exposure to our liquidity factors, we document that not one but two of the common liquidity factors are significantly related to cross-sectional differences in returns. Interestingly, neither of the two factors is related to the price dimension of liquidity, while both are related to the time and quantity dimensions. This result is robust to various model specifications. Second, and perhaps more striking, we discover that order-based liquidity measures are unable to explain return differences while common factors estimated from trade-based liquidity measures are significantly related to cross-sectional variation in realized returns. This latter result suggests that realized liquidity (as opposed to expected liquidity) may be a fundamental driver in asset returns.

**JEL Codes:** G12; G14

**Keywords:** Market microstructure; Common factor; Asset pricing; Liquidity factor; high frequency liquidity

# 1 Introduction

Our goal in this paper is to construct and test statistically-based, comprehensive liquidity measures. While the raw liquidity variables that we use are based on theory, our datasets and methodology are quite new.

A major motivation for our study is the fact that there is little knowledge about the fundamental nature of liquidity, especially at the very highest frequencies. Moreover, in spite of a large and growing literature on the relationship between asset prices and liquidity, there is currently little agreement as to what the asset pricing implications of liquidity actually are. Empirically, there is some evidence of a link between asset prices and liquidity measures such as spreads, depth and volume, as espoused by the "commonality" literature. There is, however, little concrete evidence on the existence of a time varying common liquidity factor.

In this paper, we assess whether we can strengthen the link between liquidity and asset prices by constructing statistically sound, empirical proxies for liquidity. This exercise is particularly valuable since the proxies used in the previous literature are sometimes narrow, or may not capture all the dimensions of liquidity that investors actually value, see Hodrick and Moulton (2003).

A market is said to be liquid if traders at any time during opening hours can buy or sell large numbers of shares quickly and at low costs. Harris [1990] associates this definition with four interrelated dimensions: width, depth, immediacy, and resiliency. Width measures the cost per share of liquidity, depth is the number of shares that can be traded at a given price, immediacy captures how quickly a given number of shares can be traded at a given cost, and resiliency is a measure of the ability to trade at minimal price impact.

This inherent multi-dimensionality of liquidity has led to a broad range of empirical measures. Many of the suggested liquidity measures are shown to be only modestly correlated, which has spurred some work on finding the single (or few) best measure. Our approach is different, since we combine theory and econometrics in a simple manner. Specifically, we use trade and order variables that theory suggests should capture liquidity. Then we implement dimension-reduction techniques to extract parsimonious estimates of common liquidity factors, which summarize the essential behavior of liquidity. We therefore uncover composite liquidity measures that capture as many dimensions of liquidity as possible, in a concrete statistical sense. In other words, we try to extract a few "deep" variables that capture the fundamental variations liquidity.

From a set of both order-based and trade-based liquidity measures we identify three common liquidity factors. By estimating securities' exposure to these factors, and by creating portfolios based on these exposures, we find that two of the factors significantly explain cross sectional differences in excess returns after we have controlled for each security's exposure to other traditional risk factors. Another interesting finding is that when we restrict the set of variables that are used to extract the common liquidity factors to order-based liquidity variables, we find no systematic variation in exposures and returns. On the other hand, when we restrict the set of variables to be only trade-based, there is a strong relation between exposures and realized returns.

The paper is organized as follows. Section 2 summarizes the relevant literature. In section 3 we describe the data set and the liquidity measures we use in the analysis. In section 4 we use common factor analysis to identify latent dimensions represented in different sets of liquidity variables. We then investigate the relationship between common variation in liquidity measures and asset returns in section 5. Section 6 concludes.

## 2 Literature

There is a large literature on the relationship between liquidity and asset pricing. This summary of the literature very close to that of O'Hara [2003]. Early theoretical papers find that transaction costs are irrelevant for asset prices, see Constantinides [1986], Aiyagari and Getler [1991], Heaton and Lucas [1996], Vayanos [1998], Vayanos and Vila [1998]. In general these papers argue that liquidity costs can only have a second-order effect on the level of asset prices because transactions cost are too small relative to the equilibrium risk premium to matter. The counter-argument to this, originally due to Amihud and Mendelson [1986], is that empirically asset prices do reflect liquidity costs. Several studies show a link between asset prices and a variety of liquidity measures such as spreads, depths and volumes, see Brennan and Subrahmanyam [1996] and Easley et al. [2002]. In this context liquidity is akin to a tax borne by investors. The papers most related to our work investigate the question whether liquidity can also affect the risk of holding an asset. This issue is more complex, as liquidity would then have to be time varying, or systematic in some sense. The related literature on liquidity commonality yields mixed results. Several papers find a positive relationship between returns and common variation in liquidity, including Chordia et al. [2000], Huberman and Halka [2001], Amihud (2002), and Pastor and Stambaugh [2003]. On the other hand, Hasbrouck and Seppi [2001] find little evidence of a common factor in liquidity. Therefore, a further contribution of our research is that, by constructing statistically and economically meaningful liquidity factors, we might shed additional light on the commonality debate.

To our knowledge, the only other paper using a similar statistical methodology is that of Hasbrouck and Seppi [2001]. The authors analyze liquidity measures, returns and order flows using common factor analysis. However their approach differs from ours in a few important respects. They analyze common variation across a small subset of companies for each variable. Then, they study the co-variation among liquidity measures, returns, and order flows using a canonical correlation analysis. Hence, they first study whether a part of the variation in one separate liquidity measure is common for all firms. Then, they study whether this common part is related common parts of variation in returns and order flows. Our approach is different in that we search for the shared variance across many different liquidity measures to obtain a proxy for liquidity which captures the latent dimensions of liquidity. We then investigate whether these common factors are related to cross-sectional differences in realized returns.

Holl and Winn [1995] calculate the correlation structure for 25 measures of liquidity using transactions data from the Australian Stock Exchange (ASE) in 1995. Only those measures that are similar by design are found to be correlated, indicating that different measures of liquidity capture different characteristics of assets. Using data from the Jakarta Stock Exchange (JSE), Aitken and Comerton-Fordre [2003] divide various liquidity measures into trade-based and order-based measures. Little correlation is found between the two categories. By examining changes in the liquidity measures within each category before and after an economic crisis on the JSE, the authors provide evidence that order-based measures provide a better proxy for liquidity than trade-based measures. Neither of these studies investigate shared variation in liquidity measures nor do they consider liquidity in an asset pricing perspective.

### 3 The Data

The data used in this study is provided to us by the Oslo Stock Exchange (OSE) in Norway.<sup>1</sup> In this section, we provide a brief introduction to the Norwegian stock market and the trading rules at the Oslo Stock Exchange, and describe the set of liquidity variables used in the common factor analysis in the following section.

#### 3.1 The Norwegian stock market

Norway is a member of the European Economic Area, and its equity market is among the 30 largest world equity markets by market capitalization.<sup>2</sup> Table 1 reports some general statistics on market values and trading activity for the firms listed on the OSE in the period from 1995 to 2005.

---

**Table 1** The Norwegian stock market - 1995-2005

---

The table reports some general statistics for the companies listed on the OSE in the period 1995-2005. Numbers are presented in nominal terms and in fixed 1998 prices. The nominal numbers are official statistics obtained from the web site [www.ose.no](http://www.ose.no), while the fixed prices are based on an official price index obtained from the web site [www.ssb.no](http://www.ssb.no). The table shows the number of companies listed at year-end, the market capitalization values, the number of transactions, turnover by value, and turnover velocity. The market values include all capital registered with the Norwegian Central Securities Depository (VPS). Before 1995, this only included Norwegian companies. Dividend values include dividends in companies listed at year-end. Turnover velocity is defined as the average of annualized turnover per month divided by market value at the end of each month.

Year	No of listed companies	Market value, NOK mill		No of trans.	Turnover, NOK bill		Turnover velocity
		nominal	1998 prices		nominal	1998 prices	
1995	165	289804	307648	394052	156.7	166.4	-
1996	172	389397	408601	569806	231.7	243.1	-
1997	217	556002	568509	829794	341.1	348.8	69.3
1998	235	413673	413673	846535	322.7	322.7	63.0
1999	215	582941	569835	1330674	445.6	435.6	88.6
2000	214	637856	604603	2418219	609.1	577.4	96.7
2001	212	677032	622845	2529182	566.4	521.1	86.4
2002	203	502938	456801	2047861	444.4	403.6	74.7
2003	178	689734	611466	2348086	552.5	489.8	97.7
2004	188	931650	822286	3390492	906.8	800.4	110.3
2005	219	1403267	1219172	5457734	1512.8	1314.3	128.9

---

The market has grown substantially during the last 10 years. Measured in real terms, the total market capitalization value at the end of 2005 was nearly four times the value at the end of 1995, and the turnover value in 2005 was nearly ten times the turnover value in 1995. Another notable characteristic of the market, not shown in the table, is a very high concentration of values and trading activity in a few large companies. At the end of 2005, the five largest companies (by market value and including the fixed state holdings) accounted for 61 percent of the market value of all listed firms, and around 50 percent of the total turnover value. Notable Norwegian listings include Norsk Hydro, Telenor, and Statoil.

---

<sup>1</sup>We obtained the data directly from the exchange's surveillance system. The SMARTS<sup>©</sup> system is the core of the exchange's surveillance operations. Through access to the SMARTS<sup>©</sup> database, we obtained all the information on orders and trades in the market

<sup>2</sup>Source is FIBV (International Federation of Stock exchanges).

## 3.2 Trading at the OSE

Since January 1999, the OSE has operated a fully automated computerized trading system similar to the public limit order book systems in Paris, Stockholm, and Toronto. The trading day comprises two sessions; the “pre-trade” session starting at 9:30 and ending with an opening auction at 10:00, and the “continuous trading” session from 10:00 until the trading closes at 16:00. During the pre-trade session, brokers can register trades that were executed after the close on the previous day as well as new orders. The opening auction at the end of the pre-trade session matches all registered orders at the price which maximizes the trading volume. During the continuous trading session, electronic matching of orders with crossing or equal price generates transactions. As is normal in most other electronic order driven markets, the order handling rule follows a price-time priority.<sup>3</sup>

## 3.3 Our data sample

Our data sample consists of every order and trade that occurred at the OSE in the period from February 1999 through March 2005. The order data include all order submissions, deletions and amendments of existing orders. We also know whether the order is a buy or a sell order. Thus, for each security in the data, we are able to reconstruct the full order book at any point in time. Every trade is linked to the underlying orders through an order ID. Thus, if a large order is executed against many smaller orders resulting in several smaller trades, we can trace each executed part back to the initial order.

### Data filtering

To remove very illiquid securities and securities that only have a short listing period in the data sample, we filter the sample as follows. First, each firm is required to have been listed for the entire data sample period. In addition, the firm must have been traded on at least 80 percent of the days when the Oslo Stock Exchange is open for trading (1539 days). This reduces the sample to a total of 42 securities.

To remove outliers from the reduced data sample, we check for erroneous order submissions. This is done for the largest orders submitted across all firms on each day. If we see that a large order is immediately cancelled or amended to a significantly lower volume, we correct the volume in the initial submission. In addition, we remove all odd-lot trades and orders and all trades reported as off market trades.

### Liquidity variables

To decide on a set of variables for the common factor analysis we set up a twofold goal. First, we want the variables to capture all aspects of liquidity that investors might value, based on theoretical considerations. To define these aspects, we use the four dimensions of liquidity suggested by Harris [1990]; width, depth, immediacy, and resiliency. Width measures the cost per share of liquidity, and is often proxied by the spread. Depth is the number of shares that can be traded at a certain price. Immediacy captures how quickly a given number of shares can be traded at a given cost, and resiliency is a measure of the ability to trade at minimal price impact (given non-informative trades).

---

<sup>3</sup>A new, similar trading system was introduced in the spring 2002. The reason for replacing the 1999 system was an agreement signed by OSE with the stock exchanges of Stockholm, Copenhagen and Iceland to establish a joint Nordic marketplace, known as NOREX. The NOREX exchanges are still independent entities, but the alliance has made it possible to create a joint Nordic marketplace with a common trading platform and harmonized regulations. For more information about trading on the OSE, see [www.ose.no](http://www.ose.no).

Second, we would like to investigate differences between order-based and trade-based liquidity measures, including their relative importance for investors' valuations. We therefore construct both trade- and order-based measures along the four liquidity dimensions. Before we begin a formal factor analysis, we must also ensure that the set of variables is suited to this type of analysis. One important criteria for determining the appropriateness of factor analysis is that the variable matrix have sufficient correlations. A rule of thumb is that a substantial number of correlation coefficients should be greater than 0.30.

A short description of the selected liquidity variables is given below. A more detailed description of how the measures are calculated is provided in the Appendix. The set of variables are summarized in Table 2.

- *Width.* To capture market width we include the effective spread and the relative quoted spread. The effective spread is trade-based while the quoted spread is order-based. The effective spread is calculated as two times the absolute value of the transaction price minus the bid-ask midpoint price. The relative quoted spread is the difference between the best bid and ask quotes measured as a percentage of the spread midpoint price. The effective spreads is often considered a more appropriate measure of trading costs than quoted spreads, especially for large trades.
- *Depth.* To capture market depth, we include two trade-based measures and three order-based measures. The trade-based measures are (i) trading volume, defined as the total number of shares traded during the day, and (ii) trading frequency, defined as the total number of trades across all firms during the day. The most natural order-based measure of depth is the number of shares at the inner quotes. To get a more comprehensive measure of market depth, we also include two measures based on the slope of the inner part of the order book: a "price slope", measured over the 6 first ticks and normalized over the total number of issued shares, and a "tick slope" measured in share volume, also over the 6 first ticks of the book. Both slope measures are averaged over the bid and ask sides of the book.
- *Immediacy.* To capture ability to trade quickly, we include three trade-based measures and two order-based measures. The trade-based measures are (i) the average number of seconds between trade executions and (ii) turnover, and (iii) the number of trades required for an order to be fully filled.<sup>4</sup> The order based measures are (i) fill time, defined as the number of seconds it takes to (fully) fill an order, and (ii) a measure of the order book symmetry. The order book symmetry measure is defined as the difference between the ask slope and the bid slope of the order book (over 6 ticks) divided by the added slopes. Division by the added slopes ensures that the measure is equal to 1 and -1 if one side of the book is empty.
- *Resiliency.* Resiliency is a measure of the ability to trade at a minimal price impact. Hence, it is trade-based by definition. To capture this dimension, we use a measure of how large (log) share volume is needed to move the price by 1 percent. This measure is equal to the inverse of the illiquidity ratio defined in Amihud [2002], and is calculated daily for each firm using close to close returns.

All variables are first calculated for each security on each trading date. To avoid biases due to intra-day trading patterns, we split the trading date into 6 hourly intervals. Except for the trading frequency, the share volume and the liquidity ratio, all measures are first averaged within each interval and then averaged over the 6 intervals to get a daily average. Then, a

---

<sup>4</sup>This last measure is actually a hybrid between a trade- and an order-based measure.

---

**Table 2** Liquidity measures

---

The table summarizes the liquidity proxies used in the common factor analysis. The four liquidity dimensions width, depth, immediacy, and resiliency is taken from Harris [1990]. In addition, the variables are split into trade-based and order-based measures.

---

	Width	Depth	Immediacy	Resiliency
Trade-based	Effective spread	Trading volume	Sec. between trades	Liquidity ratio
	Rel. effective spread	Trading frequency	Turnover Trades per order	
Order-based	Quoted spread	Depth, inner quotes	Order book symmetry	
	Rel. quoted spread	Price slope	Fill time	
		Tick slope		

---

cross sectional average is calculated for each date. The cross sectional average represents the market wide realization of the liquidity variable on each date. Table 3 provides some descriptive statistics for the liquidity measures.

During our sample period, the average firm value was NOK 8.33 billion and varied from a maximum of NOK 97.4 billion to a minimum of NOK 156 millions (these numbers are not shown in the table). Measured by the effective spread, the average cost of trading was NOK 0.75. Measured in relative quoted spreads, average trading costs amounted to 1.28 percent of the midpoint price. The average depth at the inner quotes for a firm was 9925 shares, and it took on average 1.6 trades to fill an order. On average, there were a total of 4236 trades in the sample firms each day, however, a median number of seconds between trades of 989 seconds (over 16 minutes) suggests that some firms in the sample were traded quite infrequently.

In the correlation matrix in Panel B of Table 3, correlations greater than 0.3 are marked gray. Visual inspection of the matrix reveals that a sufficient number of correlation coefficients are large enough to justify a factor analysis. Note also that most of the order-based and trade-based measures within the same liquidity dimensions exhibit fairly high correlations.

## 4 Common Factor Analysis

Since factor analysis is not typically used in this literature, we first discuss a few details on our research design and methodology. Factor analysis comprises a family of statistical techniques concerned with the reduction of a set of observable variables in terms of a small number of latent factors.<sup>5</sup> In this section, we first discuss the choice of factor method for extracting latent factors of liquidity as well as some other design issues. Then, we present the main results from three factor models estimated on various combinations of the liquidity variables described in section 3.

### 4.1 Some design issues

The two main factor methods are common factors and principal components. In a *common factor* analysis it is assumed that the variance can be decomposed in two: common variance

---

<sup>5</sup>For a detailed discussion of factor analysis, see for example Hair et al. [1998].

**Table 3** Descriptive statistics

The table presents descriptive statistics for the liquidity variables used in the common factor analysis. Panel A shows means, medians, standard deviations, and maximum and minimum values, while panel B shows the correlation structure (Pearson's correlation coefficients). Correlations greater than 0.3 are highlighted in gray. A description of how the liquidity variables are calculated is provided in Appendix A.

**Panel A**

ID	Variables	Mean	Median	STD	Max	Min
1	Effective spread	0.7520	0.7082	0.2170	1.5994	0.3764
2	Rel effective spread	0.97 %	0.93 %	0.27 %	2.22 %	0.49 %
3	Quoted spread	0.9880	0.9314	0.2981	1.9560	0.4940
4	Relative spread	1.28 %	1.26 %	0.37 %	3.00 %	0.62 %
5	Trading volume	9120288	7961830	5277809	42233205	808360
6	Trade frequency	4236	4151	1789	14496	485
7	Depth inner quotes	9925	9488	3432	83961	4502
8	Norm. priceslope	0.0008	0.0006	0.0006	0.0032	0.0002
9	Tick slope	6015	5710	2158	13260	2150
10	Sec. between trades	1006	989	205	1854	324
11	Turnover	0.31	0.28	0.15	1.19	0.06
12	Trades per filled order	1.60	1.61	0.09	1.92	1.25
13	Book symmetry	-0.0196	-0.0103	0.1000	0.2235	-0.3870
14	Fill time	540	472	247	1794	74
15	Liquidity ratio	0.0048	0.0045	0.0019	0.0220	0.0019

**Panel B**

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14
2	0.40													
3	0.97	0.39												
4	0.41	0.97	0.45											
5	-0.54	-0.02	-0.58	-0.08										
6	-0.28	-0.20	-0.36	-0.26	0.68									
7	-0.49	-0.14	-0.49	-0.16	0.43	0.17								
8	-0.71	-0.31	-0.76	-0.38	0.61	0.34	0.41							
9	-0.72	-0.46	-0.75	-0.51	0.68	0.45	0.59	0.73						
10	0.11	0.38	0.18	0.43	-0.29	-0.58	-0.04	-0.24	-0.30					
11	-0.31	-0.26	-0.35	-0.30	0.70	0.72	0.26	0.48	0.59	-0.52				
12	-0.42	-0.24	-0.53	-0.34	0.50	0.57	0.30	0.53	0.52	-0.49	0.48			
13	-0.04	-0.55	-0.06	-0.56	-0.11	0.11	-0.02	0.07	0.21	-0.26	0.15	0.09		
14	-0.52	0.04	-0.57	-0.03	0.42	0.09	0.29	0.61	0.36	0.06	0.14	0.46	-0.20	
15	0.39	0.67	0.38	0.66	0.08	0.08	-0.16	-0.29	-0.37	0.05	0.04	-0.11	-0.40	-0.10

that is shared by other variables in the model, and unique variance that is unique to a particular variable, including an error component. As the name suggests, a common factor analysis focus on the common variance of the observed variables. Specific variation and the error term are excluded from the analysis. In a *principal component* analysis, no distinction is made between common and unique variance. The objective is to account for the maximum portion of variance present in the original set of variables with a minimum number of composite components. Thus, while both techniques are widely used for the same purpose (data reduction), they are quite different in terms of the underlying assumptions.

We intend to identify the latent dimensions that explain why different liquidity variables



are correlated with each other. Thus, we want to extract a small number of factors to account for the intercorrelations among our observed liquidity variables. This objective calls for a common factor rather than a principal component analysis. Moreover, the scarce amount of prior knowledge we have about the composition of the variance of different liquidity measures also speaks for the common factor method.

A second design issue concerns *how many factors* to extract. While there exist several criteria or empirical guidelines for determining this number, the decision is ultimately subjective. We look at several criteria before we decide on the number of common liquidity factors. First, we use a version of the latent root criterion which says that the latent root (or eigenvalue) of the factors should exceed the average of the initial communality estimates. Second, we check that the number of factors lies close to the "elbow" of the scree plot.<sup>6</sup> We also require a minimum of three observed variables for each factor expected to emerge.

*Factor rotation* is a way to simplify the rows and columns of the factor matrix to make the factors interpretable. There are two main methods; orthogonal and oblique. In an orthogonal rotation, the axes are maintained at 90 degrees, while in an oblique rotation, there is no such restriction, meaning that the factors can be correlated with each other. The orthogonal methods are most widely used, although the oblique rotation is more flexible and also more realistic, since important underlying dimensions are not necessarily uncorrelated.<sup>7</sup> Since we want to use the factors for later regression analysis, it is most convenient for us to work with uncorrelated factors constructed using the orthogonal factor rotation.

## 4.2 Results

We estimate three different factor models. The first model is based on most of the variables described in section 3 (model A). In the second model, we include only trade-based liquidity variables (model B), and in the third model, we include only order-based liquidity variables (model C). Table 4-6 summarizes the main results from the three models. The tables report the rotated factor loadings as well as the final estimates of shared variance among the variables. We also report Kaiser's Measure of Sampling Adequacy (MSA), both overall and for the individual variables.<sup>8</sup> In general, values of MSA greater than 0.8 are considered good, while values less than 0.5 are unacceptable.

### A factor model based on all variables

Based on the criteria described above, we end up with three common factors in this model. To be able to interpret the rotated factor loadings, we must decide what constitutes a significant loading. A rule of thumb frequently used is that factor loadings greater than 0.30 in absolute value are considered to be significant. In the table, these loadings are marked in gray. Ideally, we would like to see a single significant loading for each variable on only one factor. Evidently, however, there are several split loadings in the table (variables with multiple significant loadings). Factor 1 seems to be mostly related to the width and depth dimensions, while factor 2 seems to be specially related to the trade-based measures of depth and immediacy. Factor 3 has significant loadings to the quoted spread, the order book symmetry, and the liquidity ratio.

---

<sup>6</sup>A scree plot is a plot of eigenvalues against corresponding factor numbers.

<sup>7</sup>On the other hand, the analytical procedures for performing orthogonal rotation are better developed than the procedures for oblique rotations.

<sup>8</sup>The underlying assumption of factor analysis is that there exists a number of unobserved latent variables that account for the correlations among the observed variables, such that if the latent variables were held constant, the partial correlations among the observed variables would be small. Kaiser's Measure of Sampling Adequacy is a summary measure of how small the partial correlations are relative to the ordinary correlations.

**Table 4** Results from a factor model for all liquidity variables

The table presents the main results from a factor model where the variable set includes a mixture of trade- and order-based liquidity measures along the four dimensions: width, depth, immediacy, and resiliency, as defined in Harris [1990]. MSA is the Kaiser’s measure of sampling adequacy. The factors are rotated according to an orthogonal rotation method (Varimax in SAS). Factor loadings greater than 0.30 are marked in gray.

Model A. Mixture	MSA	Shared Variance	Factor loadings		
			Factor1	Factor2	Factor3
Effective spread	0.8619	0.7089	-0.8097	-0.1052	0.2054
Relative quoted spread	0.7435	0.8006	-0.2394	-0.2454	0.8265
Trading volume	0.8087	0.8345	0.6427	0.5979	0.2530
Trade frequency	0.7877	0.7553	0.2005	0.8457	-0.0001
Depth inner quotes	0.8572	0.3829	0.6142	0.0756	-0.0008
Price slope	0.8677	0.6688	0.7504	0.2839	-0.1582
Tick slope	0.8690	0.8405	0.7938	0.3653	-0.2774
Sec. between trades	0.8356	0.5525	0.0172	-0.6858	0.2861
Turnover	0.8644	0.6998	0.3221	0.7714	-0.0320
Trades per filled order	0.8994	0.4790	0.3906	0.5545	-0.1376
Book symmetry	0.7803	0.4142	-0.0714	0.1478	-0.6223
Liquidity ratio	0.7554	0.6301	-0.3067	0.1809	0.7094
Overall MSA	0.8333				
Total shared variance/Variance explained		7.7670	3.0923	2.7910	1.8837
% of shared variance			40 %	36 %	24 %

Thus it is not easy to name this factor, although it may be related to information asymmetry. The model is quite well suited to factor analysis according to the MSA criterion. The overall MSA is 0.83, and the individual MSA numbers varies from a minimum of 0.74 for relative quoted spread to a maximum of 0.90 for the number of trades needed to fill an order.

#### A factor model on trade-based measures

In model B, we extract two common factors. Factor 1, which explains 65 percent of the shared variance among the variables, seems to be mostly related to depth and immediacy. Factor 2 is most related to spread and resiliency. Model B is acceptable according to the MSA criterion. The overall MSA is 0.71, and the individual MSA numbers varies from a minimum of 0.54 for relative effective spread to a maximum of 0.86 for the number of trades needed to fill an order.

#### A factor model on order-based measures

We also extract two common factors in Model C. The first factor explains 70 percent of the shared variance and has significant loadings from variables within all the three relevant liquidity dimensions (width, depth, and immediacy). Factor two is significantly related to the relative quoted spread, the tick slope and the order book symmetry. Model C is acceptable based on the MSA criterion. The overall MSA is 0.76, and the individual MSA numbers varies from a minimum of 0.56 for the order book symmetry to a maximum of 0.82 for the quoted spread.

**Table 5** Results from a factor analysis on trade-based liquidity measures

The table presents the main results from a factor model where the variable set includes trade-based liquidity measures along the four dimensions: width, depth, immediacy, and resiliency, as defined in Harris [1990]. MSA is the Kaiser’s measure of sampling adequacy. The factors are rotated according to an orthogonal rotation method (Varimax in SAS). Factor loadings greater than 0.30 are marked gray.

Model B. Trade-based	MSA	Shared Variance	Factor loadings	
			Factor1	Factor2
Effective spread	0.6250	0.4219	-0.4100	0.5038
Relative effective spread	0.5371	0.6834	-0.1725	0.8085
Trading volume	0.6557	0.6804	0.8245	0.0246
Trade frequency	0.8243	0.7247	0.8510	-0.0203
Sec. Between trades	0.7666	0.3785	-0.5836	0.1948
Turnover	0.7948	0.6735	0.8178	-0.0686
Trades per filled order	0.8617	0.4611	0.6434	-0.2172
Liquidity ratio	0.5845	0.6197	0.0969	0.7812
Overall MSA	0.7128			
Total variance/variance explained		4.6433	3.0347	1.6086
% of shared variance			65 %	35 %

**Table 6** Results from a factor analysis on order-based liquidity measures

The table presents the main results from a factor model where the variable set includes order-based liquidity measures along the four dimensions: width, depth, immediacy, and resiliency, as defined in Harris [1990]. MSA is the Kaiser’s measure of sampling adequacy. The factors are rotated according to an orthogonal rotation method (Varimax in SAS). Factor loadings greater than 0.30 are marked gray.

Model C. Order-based	MSA	Shared Variance	Factor loadings	
			Factor1	Factor2
Quoted spread	0.8245	0.7755	-0.8572	0.2015
Relative quoted spread	0.6906	0.6126	-0.2928	0.7258
Depth inner quotes	0.7841	0.3215	0.5637	-0.0612
Price slope	0.8049	0.7378	0.8452	-0.1532
Tick slope	0.7652	0.7717	0.7816	-0.4010
Order book symmetry	0.5579	0.4598	-0.0608	-0.6754
Fill time	0.7186	0.5146	0.6742	0.2450
Overall MSA	0.7594			
Total variance/variance explained		4.1933	2.9217	1.2716
% of total shared variance			70 %	30 %

## 5 Portfolio returns and liquidity factor exposure

As described in the previous section, the common factor analysis suggest that the joint variation in our set of liquidity variables are best captured by three orthogonal factors. Based on the factor analysis, we extract three daily score series that represent the daily realizations of the common factors. These series represent our market wide liquidity measures. In this section we

examine to what degree these liquidity measures are able to explain returns of equity portfolios in the cross-section. This will be a first step in examining whether there is a risk premium related to firms' exposure to market liquidity. These results are interesting from a practical standpoint, since sophisticated investors may actually construct and hold portfolios based on our liquidity measures.

## 5.1 Estimating the liquidity exposures

We apply a simple scheme to estimate firms' exposure to the market wide liquidity variables. By regressing the daily returns on the daily realizations of the three market wide liquidity measures for each month, we obtain three liquidity exposures for each company. These exposures will be used in the subsequent sections for portfolio creation. It is also important to note that the liquidity factors are orthogonal to each other by construction as a result of the common factor estimation.

We apply four different methods for estimating the liquidity exposures with varying adjustments for additional risk factors. In the later sections we use similar estimation models with a smaller number of liquidity factors. However, in this section we will discuss the case of three liquidity factors. Note that we estimate each firm's exposure for each month using daily data. Thus, the models described below may be viewed as the estimation for one month since we have dropped the month-year subscript on the estimated exposures. We now describe the four empirical models.

In the simplest model, we estimate each factor exposure in isolation, without considering the effects of the other two liquidity factors. Consequently, the exposures to each of the three factors,  $k \in [1, 2, 3]$ , are estimated each month in each year for each company,  $i$ , as,

$$r_{i,t} = \alpha_i^k + \lambda_i^k F_t^k + \varepsilon_{i,t} \quad (1)$$

where  $r_{i,t}$  is company  $i$ 's return on day  $t$ ,  $F_t^k$  is the realization of the market-wide factor  $k$  on date  $t$ ,  $\lambda_i^k$  is the exposure of company  $i$  to factor  $k$ . So in this framework each  $\lambda_i^k$  is estimated separately.

In the second model, we estimate each firm's exposure, taking account of the firm's exposure to the two other factors. Thus, we estimate each  $\lambda_i^k$  for each month for each company as,

$$r_{i,t} = \alpha_i + \sum_{k=1}^3 \lambda_i^k F_t^k + \varepsilon_{i,t} \quad (2)$$

In the third model, we estimate the exposure both taking account for the firm's exposure to the two other factors as well as the firm's market beta relative to the value weighted market return,

$$r_{i,t} = \alpha_i + \sum_{k=1}^3 \lambda_i^k F_t^k + \beta_i^m r_t^m + \varepsilon_{i,t} \quad (3)$$

Finally, in the fourth model, we adjust for the securities' exposure to the SMB (size) and HML (value) factors of Fama and French [1993] as well as the momentum factor (MOM) of Carhart [1997],

$$r_{i,t} = \alpha_i + \sum_{k=1}^3 \lambda_i^k F_t^k + \sum_{j=1}^4 \beta_i^j r_t^j + \varepsilon_{i,t} \quad (4)$$

where  $\beta_i^j$  denote firm  $i$ 's exposure to the market factor ( $j=1$ ), SMB factor ( $j=2$ ), HML factor ( $j=3$ ) and the momentum factor ( $j=4$ ).

In the two last models one can also consider an excess return framework where we assume that the market model (equation 3) and the Fama/French and Carhart factors (equation 4) describe the expected return of the security,  $E[r_i]$ . Then our approach assesses whether the excess return across firms have any systematic structure left to be explained by our liquidity measures, after accounting for the expected returns from these factor models. Thus, in general, we are estimating the following regressions,

$$r_{i,t} - E[r_i] = \alpha_i + \sum_{k=1}^3 \lambda_i^k F_t^k + \varepsilon_{i,t} \quad (5)$$

## 5.2 The liquidity factors

Panel A in table 7 shows the correlations of each of the liquidity factors,  $F^k$ , with the value-weighted and equally weighted market return, the Fama/French size- (SMB) and value factors (HML) and the Carhart momentum factor (MOM). All the correlations are based on daily realizations of the variables. Gray fields indicate that the Pearson correlation coefficient is significant at the 5% level. The first thing to note from the table is that all the correlations are low. In addition, we see that all the three factors have a significant correlation with the market return. Furthermore, we see that the two first factors are positively correlated with market return, while the third factor is negatively correlated with the market. Panel B in table 7 shows the monthly returns and standard deviations of the same variables. The returns and standard deviations of the three liquidity factors are not included in the table since they all are normalized to have a zero mean and unit standard deviation.

---

**Table 7** Correlation of liquidity factors with Fama-French and Carhart factors

---

The table show the daily correlations between the three liquidity factors from the common factor analysis with the market return (equally weighted and value weighted), the Fama/French factors and the Carhart momentum factor (MOM). Gray fields indicate that the Pearson correlation coefficient is significant at the 5% level.

**(a) Factor correlations**

	Factor 1	Factor 2	Factor 3
VW market return	0.071	0.082	-0.041
EW market return	0.109	0.100	-0.053
SMB factor	-0.048	-0.050	0.011
HML factor	-0.025	-0.035	-0.005
MOM factor	-0.008	0.040	0.065

**(b) Monthly factor returns**

	Mean	Std.dev
VW market return	2.33 %	5.0 %
EW market return	2.76 %	3.7 %
SMB factor	2.47 %	5.7 %
HML factor	1.17 %	5.7 %
MOM factor	1.25 %	5.6 %

---

## 5.3 Monthly returns on liquidity portfolios

The first question we examine is whether firms' exposure to variations in our market-wide liquidity variables is related to their realized returns. We start by estimating the liquidity exposure for each firm in the sample for each month based on the four models in section 5.1. For each month we then sort all firms based on their exposure to each factor and assign each firm to a quartile based on their estimated exposure. The firms in the lowest quartile, with the

lowest (closest to zero or most negative) exposure, will be assigned to portfolio 1 and so on. Since we have three factors, this will result in three sortings (one for each factor) for each of the four models.<sup>9</sup> Finally, we calculate the average realized return for the portfolios resulting from sorting the companies relative to their exposure to each of the three factors and for each exposure estimation model.

In table 8 we show the results from this analysis. The four panels in the table refer to the different methods that are used to estimate firms' exposure to the three liquidity factors. Recall that in model 1 we estimate each exposure in isolation, in model 2 we estimate all liquidity exposures simultaneously, in model 3 we correct for market risk, and in model 4 we correct for the exposure to the market, the Fama/French factors and the Carhart factor.

Each panel of the table shows the average monthly return for the firms in each portfolio for each factor sorting. The last two lines in each panel shows the difference in returns between the 4th and the 1st portfolio and the t-value from a test that the returns of portfolio 1 is equal to portfolio 4. If none of our liquidity factors contain information about realized returns, we should see no patterns across portfolios for any of the factor sortings. While this is the case for the first factor, there is, on the other hand, strong evidence that both factor 2 and factor 3 are important for explaining returns across portfolios. Importantly, this effect is robust to the situation depicted in panel (d), when we control for other risk factors proposed in the literature. As we move from panel (a) to panel (d) we see that the additional risk variables explain some of the differences in portfolio returns. Importantly, we obtain a monotonic pattern in the returns, which indicates a very consistent pricing effect of our liquidity on returns. Even after correcting for the market return, the Fama/French factors and the Carhart factor the difference in returns between portfolio 1 and 4 is 3.7% and 4.6% per month when we base the portfolio construction on factor 2 and factor 3 respectively. These results are very striking, since they suggest that not one but two liquidity factors could be responsible for the observed pattern of returns.

**Table 8** Return on liquidity portfolios

The tables shows average portfolio returns for portfolios created based on the underlying firms' exposures the three liquidity factors obtained from the common factor analysis. The two last lines of each panel shows the difference in average monthly returns between portfolio 4 and portfolio 1 and the t-statistics from a test of a hypothesis that the two portfolios have the same average return. Panel (a) shows the results when exposures are estimated as in eq. 1, panel (b) shows the results when exposures are estimated as in eq. 2, panel (c) shows the results when exposures are estimated as in eq. 3 and panel (d) shows the results when exposures are estimated as in eq. 4.

(a) Model 1				(b) Model 2			
Portfolio	Factor 1	Factor 2	Factor 3	Portfolio	Factor 1	Factor 2	Factor 3
1	2.22 %	-2.00 %	-1.97 %	1	1.81 %	-1.18 %	-1.18 %
2	1.04 %	0.29 %	0.22 %	2	0.86 %	0.08 %	-0.22 %
3	1.00 %	2.45 %	1.87 %	3	1.09 %	2.06 %	2.60 %
4	0.93 %	4.63 %	5.27 %	4	1.43 %	4.37 %	4.13 %
Diff. 4-1	-1.29 %	6.63 %	7.23 %	Diff. 4-1	-0.38 %	5.55 %	5.32 %
t-test diff.	-1.10	7.52	6.64	t-test diff.	-0.34	5.47	5.54

(c) Model 3				(d) Model 4			
Portfolio	Factor 1	Factor 2	Factor 3	Portfolio	Factor 1	Factor 2	Factor 3
1	1.89 %	-1.55 %	-0.77 %	1	1.66 %	-0.64 %	-0.57 %
2	0.51 %	0.51 %	0.33 %	2	0.35 %	0.68 %	0.36 %
3	1.06 %	2.30 %	2.11 %	3	1.04 %	2.18 %	1.48 %
4	1.74 %	4.06 %	3.62 %	4	2.16 %	3.04 %	4.04 %
Diff. 4-1	-0.16 %	5.61 %	4.39 %	Diff. 4-1	-0.51 %	3.67 %	4.61 %
t-test diff.	-0.10	5.21	5.30	t-test diff.	0.55	4.21	5.23

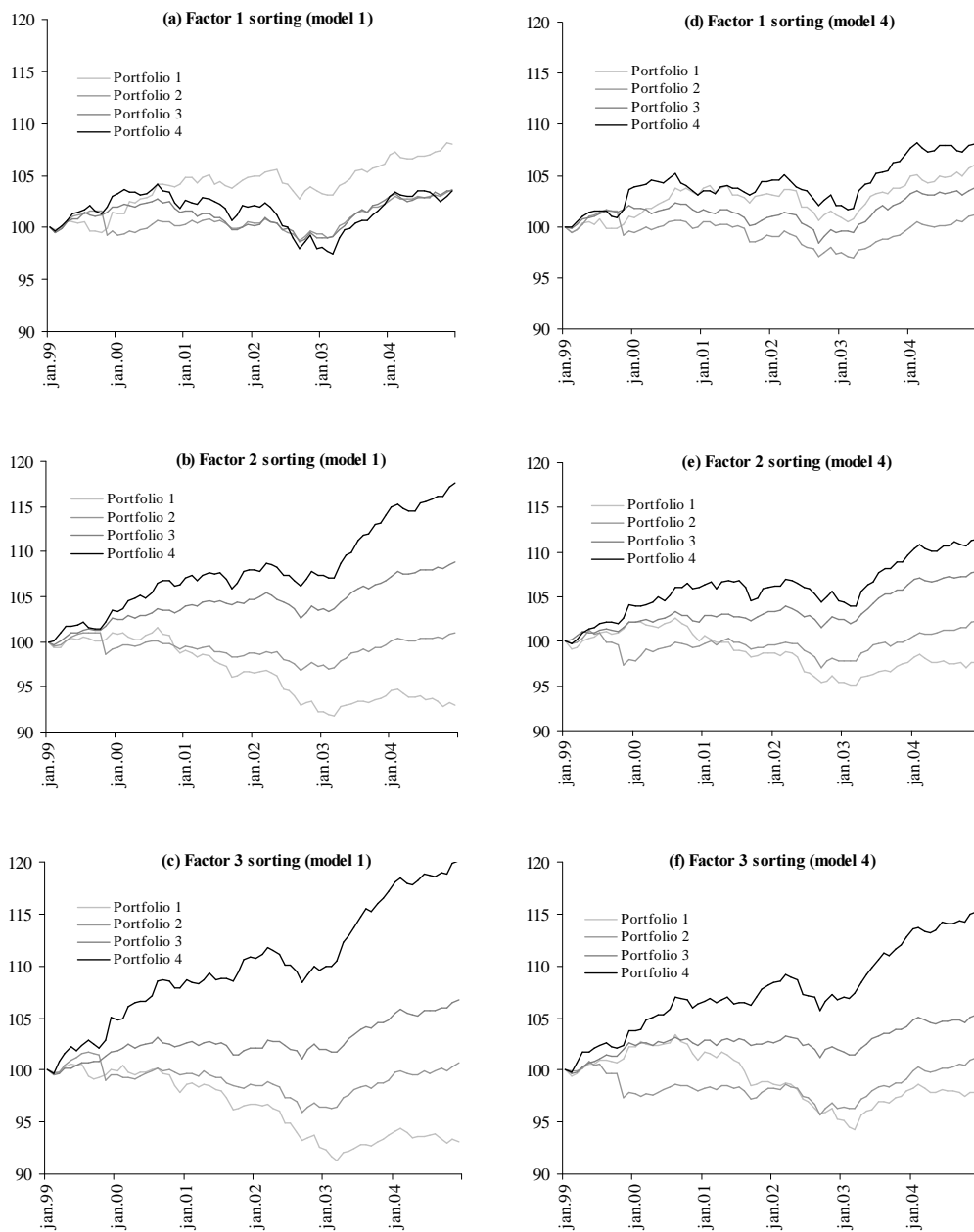
In figure 1 we visualize the accumulated portfolio values when we start with a value of 100

<sup>9</sup>So far we have only examined sortings based on each variable separately, not based on all three factors simultaneously.

in each portfolio at the beginning of our sample period and change the composition of securities in each portfolio every month. In plots (a) to (c) we show the portfolio values when we create portfolios based on the exposures to each of the three factors estimated from model 1 (with no additional corrections). In plots (d) to (f) we show the similar portfolios for each factor when we have estimated the exposures from model 4. From these figures we see the same pattern as in table 8. First, in figure (a) and (d) there is no systematic difference in portfolio returns for portfolios sorted on factor 1, but there is a large difference when we sort on factor 2 in figure (b) and (e) and factor 3 (c) and (f). Second, we see that the correction for additional risk factors reduces the difference in portfolio returns. However, the portfolio values still increases monotonically as we move from portfolio 1 to portfolio 4. Hence the robustness of this monotonic pattern is strongly suggestive of a *dual* asset pricing role for liquidity.

**Figure 1** Long run performance of liquidity factor portfolios

Plots (a) to (c) shows the cumulative portfolio values for the four portfolios when we sort portfolios each month on each of the liquidity factors based on the exposures estimated as in eq.1. Plots (d) to (f) shows the cumulative portfolio values for the portfolios based on the exposure estimates from eq.4 where we control for the exposure to the market portfolio Fama/French factors and the Carhart momentum factor.



#### 5.4 Difference between order- and trade-based liquidity factors

As discussed in the common factor section, recent literature on how to measure liquidity suggests that one should look more carefully at order-based (ex-ante) measures as opposed to traditionally applied trade-based (ex-post) liquidity measures. We therefore conduct a similar analysis to that of the previous section, with liquidity factors extracted separately from trade-based and order-based liquidity variables.



## Trade-based liquidity factors

When we restrict the liquidity variables in the common factor analysis to be trade-based, we obtain two common factors. In the following analysis we use these two factors as our liquidity measures when we do a similar analysis to that of section 5.3.

In table 9 we see the similar patterns as earlier for both liquidity factors and across all model specifications. However, in the panel (d) the difference in returns between the 1st and 4th portfolio is not significantly different when we sort companies based on their exposure to factor 1. However, the difference in returns between firms in the groups remains highly significant when we base the portfolio sorting on factor 2. These results can also be seen in figure 2 where we see that the difference in portfolio returns is reduced dramatically from figure 2a (with no correction for additional factors) to figure 2c (where we correct for all the additional risk factors).

---

**Table 9** Return on liquidity portfolios - trade based factors

---

The tables shows average portfolio returns for portfolios created based on the underlying firms' exposures the two liquidity factors obtained from the common factor analysis when we only consider trade-based liquidity variables. The two last lines of each panel shows the difference in average monthly returns between portfolio 4 and portfolio 1 and the t-statistics from a test of a hypothesis that the two portfolios have the same average return. Panel (a) shows the results when exposures are estimated as in eq. 1, panel (b) shows the results when exposures are estimated as in eq. 2, panel (c) shows the results when exposures are estimated as in eq. 3 and panel (d) shows the results when exposures are estimated as in eq. 4.

<b>(a) Model 1</b>			<b>(b) Model 2</b>		
Portfolio	Factor 1	Factor 2	Portfolio	Factor 1	Factor 2
1	-1.89 %	-2.02 %	1	-0.45 %	-1.32 %
2	0.06 %	-0.06 %	2	0.67 %	0.42 %
3	2.63 %	2.06 %	3	1.88 %	1.19 %
4	4.56 %	5.43 %	4	3.15 %	5.07 %
Diff. 4-1	6.45 %	7.45 %	Diff. 4-1	3.59 %	6.40 %
t-test diff.	6.78	7.84	t-test diff.	3.47	6.85

<b>(c) Model 3</b>			<b>(d) Model 4</b>		
Portfolio	Factor 1	Factor 2	Portfolio	Factor 1	Factor 2
1	-0.33 %	-1.06 %	1	0.26 %	-0.54 %
2	0.80 %	0.73 %	2	0.20 %	0.95 %
3	1.90 %	1.54 %	3	2.61 %	1.49 %
4	2.87 %	4.10 %	4	2.15 %	3.37 %
Diff. 4-1	3.21 %	5.16 %	Diff. 4-1	1.88 %	3.91 %
t-test diff.	2.94	5.94	t-test diff.	1.95	4.12

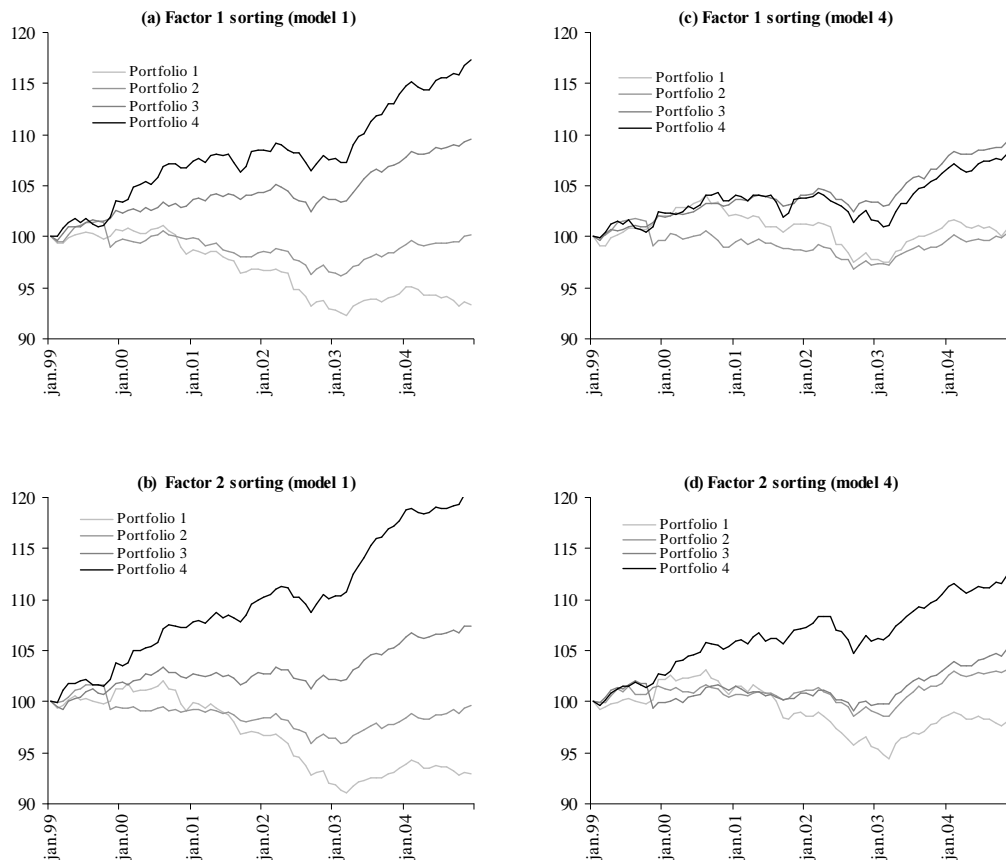
---

---

**Figure 2** Liquidity factor portfolios - trade based liquidity factors

---

Plots (a) and (b) shows the cumulative portfolio values for the four portfolios when we sort portfolios each month on each of the liquidity factors based on the exposures estimated as in eq.1 relative to common factors extracted from a set of trade-based liquidity variables. Plots (c) to (d) shows the cumulative portfolio values for the portfolios based on the exposure estimates from eq.4 where we control for the exposure to the market portfolio Fama/French factors and the Carhart momentum factor.



---

**Order-based liquidity factors**

When we restrict the liquidity variables in the common factor analysis to be order-based, we also obtain two common liquidity factors. Table 10 shows the results from this analysis. It is interesting to note that there is no systematic difference in portfolio returns when we construct the common liquidity factors only from order-based liquidity variables. The only visible pattern in the tables are that there seem to be a u-shaped pattern across portfolio returns. At this point we do not have enough information to speculate whether that pattern has any economic meaning. The lack of any significant systematic patterns across portfolios in table 10 is also visualized in figure 3.

Although this last result is disappointing from a market microstructure perspective, it is still very interesting with respect to how liquidity is best measured. This result is of course not any evidence against order-based measures being important for measuring liquidity. On the other hand, assuming that we have used reasonable and appropriate order-based liquidity variables as inputs to our common factor analysis, we can conclude that it seem to be the common variation in trade based liquidity variables that best explain cross sectional differences in returns.

---

**Table 10** Return on liquidity portfolios - order based factors

---

The tables shows average portfolio returns for portfolios created based on the underlying firms' exposures the two liquidity factors obtained from the common factor analysis when we only consider order-based liquidity variables. The two last lines of each panel shows the difference in average monthly returns between portfolio 4 and portfolio 1 and the t-statistics from a test of a hypothesis that the two portfolios have the same average return. Panel (a) shows the results when exposures are estimated as in eq. 1, panel (b) shows the results when exposures are estimated as in eq. 2, panel (c) shows the results when exposures are estimated as in eq. 3 and panel (d) shows the results when exposures are estimated as in eq. 4.

Portfolio	<b>(a) Model 1</b>		Liquidity exposure	<b>(b) Model 2</b>	
	Factor 1	Factor 2		Factor 1	Factor 2
1	1.42 %	0.92 %	1	1.26 %	1.41 %
2	0.72 %	0.69 %	2	0.59 %	0.19 %
3	0.90 %	1.82 %	3	1.44 %	1.41 %
4	2.17 %	1.77 %	4	1.91 %	2.20 %
Diff. 4-1	0.75 %	0.85 %	Diff. 4-1	0.65 %	0.79 %
t-test diff.	0.94	-1.13	t-test diff.	0.81	-0.20

Portfolio	<b>(c) Model 3</b>		Liquidity exposure	<b>(d) Model 4</b>	
	Factor 1	Factor 2		Factor 1	Factor 2
1	1.97 %	1.41 %	1	1.66 %	1.51 %
2	0.43 %	0.19 %	2	0.67 %	0.14 %
3	0.97 %	1.41 %	3	1.71 %	1.51 %
4	1.84 %	2.20 %	4	1.16 %	2.05 %
Diff. 4-1	-0.14 %	0.79 %	Diff. 4-1	-0.50 %	0.54 %
t-test diff.	-0.04	-0.04	t-test diff.	-0.21	-0.39

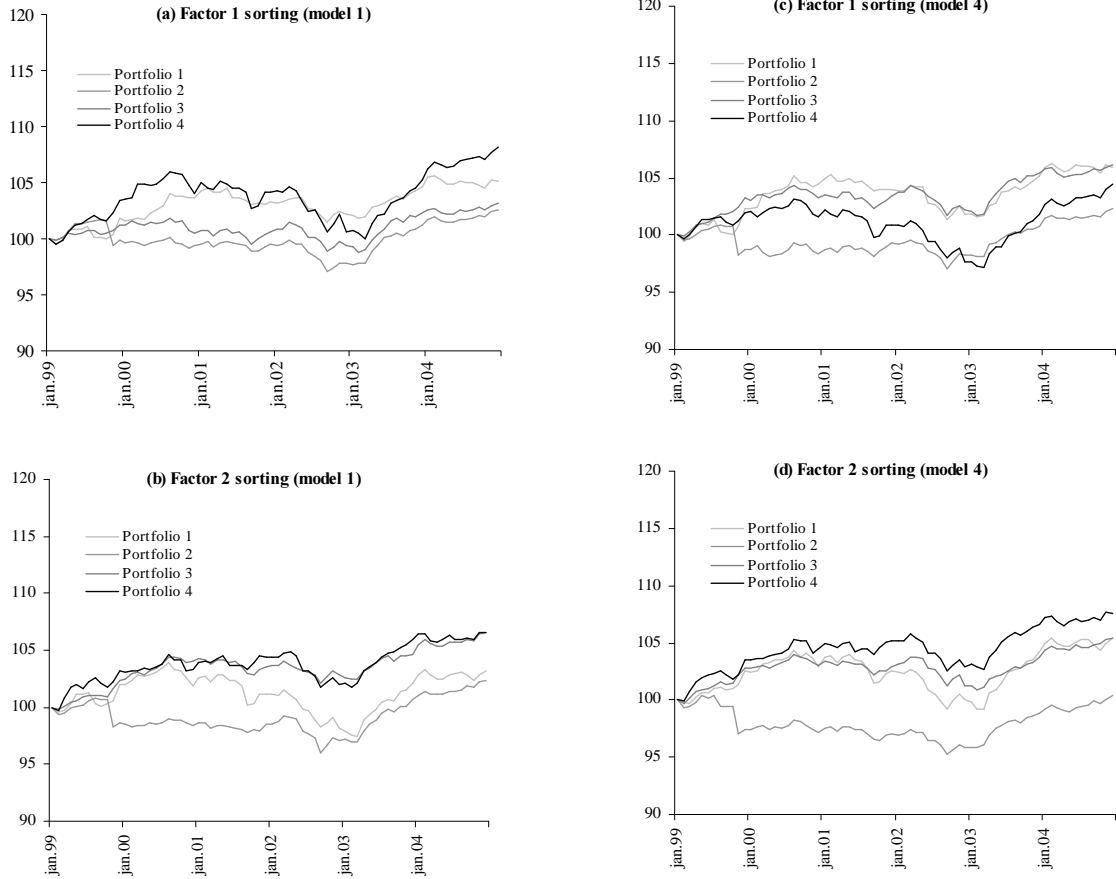
---

---

**Figure 3** Liquidity factor portfolios - order based liquidity factors

---

Plots (a) and (b) shows the cumulative portfolio values for the four portfolios when we sort portfolios each month on each of the liquidity factors based on the exposures estimated as in eq.1 relative to common factors extracted from a set of order-based liquidity variables. Plots (c) to (d) shows the cumulative portfolio values for the portfolios based on the exposure estimates from eq.4 where we control for the exposure to the market portfolio Fama/French factors and the Carhart momentum factor.



---

## 6 Conclusion

There is currently little research on how best to define and test liquidity, based on theoretical and statistical reasons. This makes it difficult to agree on the asset pricing implications of liquidity, since many of the liquidity measures proposed in the literature are only modestly correlated with each other. In this paper we utilize a common factor approach in order to extract fundamental, common liquidity factors from a large set of liquidity variables. In a sample of several liquidity variables from an electronic limit order market we find that there are three common factors. We document two main findings. First, we discover that not one but two of these factors are significantly related to cross-sectional variation in realized returns. This suggests that there might be a *dual* role of liquidity in asset pricing. Interestingly, the two factors reflect time and quantity aspects of liquidity, instead of price aspects.

Second, we find that common factors constructed from only order-based measures is not related to the cross section of returns, while trade-based measures are strongly related to returns. This result indicates that realized liquidity (rather than expected liquidity) is a fundamental driver in asset returns.

In sum, our results are interesting and important, since the portfolios that we form can be easily utilized by moderately sophisticated investors. Consequently, investors may (in principle) be able to hedge specific types of liquidity risk by holding one or both of our factors.

The current version of the paper is planned to be extended in several directions. First of all, in the current version of the paper we do not conduct any asset pricing test of whether our estimated liquidity factors are actually priced. This is something that will be incorporated in future versions. Secondly, more work is needed to interpret the liquidity factors that are estimated from the common factor analysis. Finally, we are also in the process of applying a similar analysis to liquidity variables from the US as a robustness check.

## A Calculation of liquidity variables

The appendix describes how the various liquidity measures used in the common factor estimation are calculated. To make the data discrete and have a common time frame we use hourly windows, starting from 10:00 am until 16:00 pm. Thus, we have 6 one hour intervals during each trading day. If not otherwise stated, the measures are first average within each interval, and then averaged over these intervals to get a daily average. For simplicity, summing and averaging operators as well as security and time indicators are suppressed in the equations. The variables are presented in alphabetical order.

**Depth at the inner quotes** is defined as the average share volume at the best quotes. The measure is calculated as

$$\frac{v^a + v^b}{2}$$

where  $v^a$  is the share-volume at the best ask price and  $v^b$  is the share-volume at the best bid price.

**Effective spread** is calculated for each trade as

$$2|p^k - \frac{p^a + p^b}{2}|$$

where  $p^k$  is the trade price, and  $p^b$  and  $p^a$  are respectively the best bid and ask prices at the time of the trade.

**Fill time** is defined as the number of seconds it takes to (fully) fill an order.

**Liquidity ratio** measures how large (log) share-volume is needed to move the price by 1%, i.e, it is the inverse of the illiquidity measure proposed by Amihud [2002]. The measure is calculated per day for each firm, using close to close returns as

$$\frac{\log(V)}{|r| * 100}$$

where  $V$  is the total daily share volume, and  $r$  is the open to close return.

**Normalized price slope** is the Price slope normalized relative to the number of issued shares in the firm, i.e

$$\text{Price slope}/NOSH$$

where **Price slope** is defined below and **NOSH** is the number of issued shares in the firm.

**Order book symmetry** measures the symmetry of the limit order book. It is calculated as the difference between the ask and bid slopes over the first 6 ticks of the order book divided by the added 6 tick slopes (to ensure that the measure is equal to 1 and -1 if one side in the order book is empty), i.e

$$\left(\frac{v_6^a - v_0^a}{p_6^a - p_0^a} - \frac{v_6^b - v_0^b}{p_0^b - p_6^b}\right) / \left(\frac{v_6^a - v_0^a}{p_6^a - p_0^a} + \frac{v_6^b - v_0^b}{p_0^b - p_6^b}\right)$$

where  $v_0^b$  is the share volume at the best bid quote,  $v_6^b$  is the cumulative share volume 6 ticks away from the best bid quote,  $v_0^a$  is the share volume at the best ask quote,  $v_6^a$  is the cumulative

share volume 6 ticks away from the best ask quote,  $p_0^b$  and  $p_0^a$  are respectively the best bid and ask quotes, and  $p_6^b$  and  $p_6^a$  are respectively the bid and ask quotes 6 ticks away from the best quotes.

**Price slope** is defined as the average of bid and ask slopes over the first 6 ticks of the order book computed relative to the price levels in the book, i.e

$$\left\{ \frac{v_6^a - v_0^a}{p_6^a - p_0^a} + \frac{v_6^b - v_0^b}{p_0^b - p_6^b} \right\} / 2$$

where  $v_0^b$  is the share volume at the best bid quote,  $v_6^b$  is the cumulative share volume 6 ticks away from the best bid quote,  $v_0^a$  is the share volume at the best ask quote,  $v_6^a$  is the cumulative share volume 6 ticks away from the best ask quote,  $p_0^b$  and  $p_0^a$  are respectively the best bid and ask quotes, and  $p_6^b$  and  $p_6^a$  are respectively the bid and ask quotes 6 ticks away from the best quotes.

**Quoted spread** is the difference between the best ask and bid quotes,

$$p^a - p^b$$

where  $p^a$  is the best ask quote and  $p^b$  is the best bid quote.

**Relative effective spread** is calculated as the effective spread divided by the bid-ask midpoint price,

$$\frac{2|p^k - \frac{p^a + p^b}{2}|}{(p^a + p^b)/2}$$

where  $p^k$  is the trade price,  $p^a$  is the best ask price, and  $p^b$  is the best bid price.

**Relative quoted spread** is calculated as the quoted spread divided by the bid-ask midpoint price,

$$\frac{p^a - p^b}{(p^a + p^b)/2}$$

where  $p^a$  is the best ask price, and  $p^b$  is the best bid price.

**Seconds between trades** is defined as the average number of seconds between trade executions.

**Tick slope** is defined as the average slope of the bid and ask side of the order book, where the slopes are measured over the 6 first ticks of the book and computed relative to the number of ticks away from the best quotes,

$$\left\{ \frac{v_6^a - v_0^a}{6} + \frac{v_6^b - v_0^b}{6} \right\} / 2$$

where  $v_0^b$  is the share volume at the best bid quote,  $v_6^b$  is the cumulative share volume 6 ticks away from the best bid quote,  $v_0^a$  is the share volume at the best ask quote,  $v_6^a$  is the cumulative share volume 6 ticks away from the best ask quote.

**Trades per order** is calculated by tracking each order from its initial submission into the system (accounting for subsequent amendments) and counting the number of trades that are required for the order to be fully filled.

**Trading frequency** is measured as the total number of trade executions across all firms during the day.

**Trading volume** is measured as the total number of shares traded in the security during the day.

**Turnover** is measured as the trading volume divided by the number of issued shares in the company.



## References

- Michael Aitken and Carole Comerton-Fordre. How should liquidity be measured? *Pacific-Basin Finance Journal*, 11:45–59, 2003.
- S. Rao Aiyagari and Mark Getler. Asset returns and transactions cost and uninsured individual risk: A stage iii exercise. *Journal of Monetary Economics*, 27:309–331, 1991.
- Y Amihud and H Mendelson. Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17:223–249, 1986.
- Yakov Amihud. Illiquidity and stock returns: cross section and time series effects. 5:31–56, 2002.
- Michael Brennan and Avanidhar Subrahmanyam. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics*, 41:441–464, 1996.
- Mark M. Carhart. On persistence in mutual fund performance. *Journal of Finance*, 52:57–82, 1997.
- Tarun Chordia, Richard Roll, and Avanidhar Subrahmanyam. Commonality in liquidity. *Journal of Financial Economics*, 56:3–28, 2000.
- George Constantinides. Capital market equilibrium with transaction costs. *Journal of Political Economy*, 94:842–862, 1986.
- David Easley, Soeren Hvidkjaer, and Maureen O’Hara. Is information risk a determinant of asset returns? *Journal of Finance*, 57:2185–2222, 2002.
- E. F. Fama and K. R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33:3–56, 1993.
- Joseph F. Hair, Rolph E. Anderson, Ronald L. Tatham, and William C. Black. *Multivariate Data Analysis*. Prentice-Hall International, INC, 1998.
- Lawrence Harris. Liquidity, trading rules, and electronic trading systems. *New York University Salomon Center Monograph Series in Finance and Economics*, 4, 1990.
- Joel Hasbrouck and Duane J. Seppi. Common factors in prices, order flows, and liquidity. *Journal of Financial Economics*, pages 383–411, 2001.
- John Heaton and Deborah Lucas. Evaluating the effects of incomplete markets on risk sharing and asset prices. *Journal of Political Economy*, pages 443–487, 1996.
- Thorsten Holl and Roland Winn. Comparability of different measures of liquidity on the Australian stock exchange. Working Paper, Security Industry Research Center of Asia-Pacific, 1995.
- Gur Huberman and Dominic Halka. Systematic liquidity. *Journal of Financial Research*, pages 161–178, 2001.
- Maureen O’Hara. Presidential address: Liquidity and price discovery. *Journal of Finance*, pages 1335–1354, 2003.

Lubos Pastor and Robert Stambaugh. Liquidity risk and expected stock returns. *Journal of Political Economy*, pages 642–685, 2003.

Dimitri Vayanos. Transactions costs and asset prices: a dynamic equilibrium model. *Review of Financial Studies*, pages 1–58, 1998.

Dimitri Vayanos and Jean-Luc Vila. Equilibrium interest rates and liquidity premium with transactions costs. *Review of Financial Studies*, pages 1–58, 1998.