Systematic Liquidity in the Xetra Order Book: A Multi-Stage Approach

Emanuel Kopp*, Michael Hütl, Otto Loistl, Johannes Prix

Vienna University of Economics and Business Administration, Austria

Abstract

The latest developments in the literature on market-wide liquidity are the investigation of order-driven market structures, the application of higher data frequencies, and there is also a shift towards a demand/supply perspective of liquidity. But most existing studies concentrate exclusively on liquidity around the spread, which represents only a small area of the liquidity provided by limit orders in the order book. We apply liquidity measures that capture *different non-overlapping tradability aspects of liquidity* in the entire limit order book. Since conventional PCA methods can be strongly affected by the presence of outliers in the sample, we rely on a robust principal component analysis method based on the Projection-Pursuit principle (Huber (1985)) to estimate the systematic liquidity components. Moreover, a PCA methodology allows no economic interpretation of the systematic factors. Therefore we propose a multi-stage PCA and regression approach that allows a more detailed investigation of cross-sectional liquidity determinants and their interactions. Additionally, we apply several other trading-related measures that allow to capture information from the (entire) limit order book, and study the relation of their market-wide factors to systematic factors in liquidity. This is the first empirical study on non-idiosyncratic liquidity components that investigates *different time periods during the trading session* based on complete tick-by-tick order book data from the Xetra trading system.

JEL Classification: G15 – International Financial Markets
G24 – Investment Banking

KEY WORDS: financial markets, market microstructure, commonality in liquidity, robust principal components analysis, order book liquidity measurement, intraday trading

**Correspondence Address:* Mag. Emanuel Kopp, Vienna University of Economics and Business Administration, Investment Banking and Catallactics Group, Althanstrasse 39-45, 1090 Vienna, Austria. Tel.: +43-1-31336-4173; Fax: +43-1-31336-761. Correspondence-Emails: emanuel.albin.kopp@wu-wien.ac.at (Mag. Emanuel Kopp), michael.huetl@wu-wien.ac.at (Mag. Michael Hütl), otto.loistl@wu-wien.ac.at (o.Univ.-Prof. Dipl.-Kfm. Dr. Dr.h.c. Otto Loistl), johannes.prix@wu-wien.ac.at (Mag. Johannes Prix).

Contents

1 Introduction

The structure of established asset pricing models indicates that the market-wide context of liquidity has been disregarded for many years. But the topic has regained attention. Today it is well documented in the academic literature that the liquidity of individual stocks in a market is considerably determined by latent market-wide factors. This strong evidence of *commonality in liquidity* has far-reaching consequences for financial theory, exchange organizations, regulators, and individual market participants. For the individual trader, commonality in liquidity means that some part of liquidity risk is systematic and, hence, cannot be diversified. Consequently, conventional asset pricing models that do not consider the second moment of liquidity fail to correctly assess liquidity risk. Finally, the understanding of these systematic liquidity factors is essential for effectively regulating and stabilizing the financial system.

Detailed intraday order book data has not been available until recently. In particular the early studies on market-wide liquidity usually applied price or quote data in monthly, weekly, or daily frequency obtained from the *ISSM Transactions File Database* or the *NYSE Trade Quotation (TAQ) Database*. Since traders are most severely exposed to changes in market-wide liquidity risk if there is a need to trade high volumes during the trading session, an investigation of market-wide liquidity in the sense of *tradability aspects* should also be studied by the use of *intraday data*.

One shortcoming of most existing papers is the almost exclusive concentration on liquidity around the spread. But the liquidity at best bid and best ask is only a small part of the liquidity provided by limit orders in the order book. For instance, Kempf and Mayston (2006) show that more than twenty percent of all transactions in the Xetra system walk up the order book. Hence, there is also valuable trading information about liquidity beyond best bid and best ask. To adress this shortcoming, we apply different *non-overlapping tradability measures of liquidity* that allow the investigation of *liquidity in different areas* of the order book.

Our empirical analysis is based on high-frequency order book data from the *Xetra* trading system. The dataset consists of each tick-by-tick Xetra order book entry for all DAX-30 stocks during the period 6-13 December 2006. Xetra operates as an open limit order book. Compared to the Paris stock exchange, not only the best five quotes are displayed to the traders. In contrast to the NYSE, every trader has insight into the complete order book[1]. Each trader can easily compute how far his or her order would have to walk up the book to get fully executed. Hence, our data set is especially appropriate for an *intraday* investigation of liquidity commonality in the *entire* limit order book.

First, we address the question whether also market-wide factors determine the liquidity of individual stocks. For this purpose, we conduct cross-sectional principal components analysis on the liquidity measures applied in this study. Since the results obtained from conventional PCA approaches can be strongly affected by the presence of outliers in the sample, we apply a much more robust method that arises from the *Projection-Pursuit* principle (Huber (1985)). Unfortunately, a PCA methodology only allows to examine if common factors *exist*, but offers no economic (or causal) interpretation of the extracted latent determinants. To get to an interpretation of the systematic liquidity components, we propose a multi-stage

[1] An exception are hidden orders (iceberg orders).

PCA and regression approach using non-overlapping cost-of-roundtrip measures. To examine whether also the systematic factors of different liquidity variables are correlated cross-sectionally, we run canonical correlations analysis between the principal components of the different variables in the cross-section. In addition, we apply several other trading-related measures and study the relation between their market-wide determinants and the systematic liquidity factors. To investigate how commonality behaves during the trading session, we analyze three different time periods during continuous trading. Detailed regression results between the systematic factors of the different PCA-stages are also reported in this paper. From regression analysis we additionally find out which determinants count for the intraday-variation of the different measure-independent systematic liquidity components.

The paper is organized as follows: Section 2 discusses the existence and consequences of systematic market-wide liquidity factors and briefly surveys the literature on market-wide liquidity. Section 3 informs about the Xetra trading system of the Frankfurt Stock Exchange (FSE) and reports the structure of our data-set. The description of the liquidity measures and the Projection-Pursuit PCA applied can be found in section 4. In section 5, we conduct several principal components and canonical correlation analysis between the systematic factors extracted from different measures. In section 6, we analyze whether also measure-independent systematic liquidity factors exist. Based on measure-specific systematic factors in liquidity, we extract *global* systematic liquidity factors and study *factor interactions* across different PCA stages. Section 7 summarizes the most important results of this study and concludes.

2 Market-Wide Liquidity

2.1 Literature Overview and Recent Developments

Until the end of the 1990ies, researchers focused almost exclusively on the liquidity of individual stocks.[2] Hasbrouck and Seppi (2001) state that "until recently, however, little direct empirical research has been conducted on the magnitudes of cross-sectional interactions at the microstructure level. (...) This focus on stocks in isolation (...) left us ignorant of even the most basic facts about cross-sectional interactions between stocks."[3]

Probably the most prominent empirical papers that investigate the time-variation of market-wide liquidity are Chordia et al. (2000) and Hasbrouck and Seppi (2001). For instance, Chordia et al. (2000) apply different spread measures and quoted depths to measure liquidity on NYSE listed stocks during 1992, and show that individual liquidity measures of different stocks co-move with each other. They argue that *commonality in liquidity* empirically manifests in time-series co-movement in liquidity because of latent common determinants across stocks, and apply a market model that relates individual stock liquidity to market liquidity, which is similar to the CAPM, but for liquidity. Concretely, they run regressions to measure the sensitivity of individual stock liquidity to market liquidity, which they define as the average of all stocks' liquidity (see Chordia et al. (2000)). They conclude that inventory

[2]For instance, O'Hara (1995) provides a survey on the literature on liquidity from a single-asset perspective.
[3]Hasbrouck and Seppi (2001), p. 384.

effects[4] and asymmetrically distributed information might be the sources of market-wide liquidity co-movements.

Different from the market-model approach, Hasbrouck and Seppi (2001) rely on principal components analysis (PCA) of different liquidity measures separately to extract latent, market-wide factors that systematically drive the liquidity of individual stocks. In recent years, the use of PCA has become the most popular methodology to study cross-sectional liquidity commonality. Hasbrouck and Seppi (2001) apply bid-ask spreads, depths, and quote-slope measures to the thirty Dow-Jones stocks during 252 trading days in 1994. While Chordia et al. (2000) report that the extracted liquidity determinants explain more than thirty percent of daily changes in liquidity, the results of Hasbrouck and Seppi (2001) are less supporting. They do not detect significant evidence of commonality in the applied liquidity measures. Conversely, they find strong evidence for the existence of market-wide factors in order-flows and stock returns.

The latest developments in the literature on market-wide liquidity factors are the investigation of different aspects of liquidity, a shift towards a demand/supply perspective of liquidity, and the use of higher data frequency. The most recent development in the commonality literature is the investigation of order-driven market structures (see Zheng and Zhang (2006)). For instance, Domowitz and Wang (2002) investigate an order data-set of the Australian ASX during 2000 and argue that liquidity commonality is due to supply and demand co-movements. They measure liquidity as a functional of supply and demand schedules, and liquidity commonality is measured by functional covariance. The liquidity measure they apply is the gap, or distance, between a stock's supply and demand schedules.

Korajczyk and Sadka (2007) estimate a measure of systematic liquidity risk across a set of eight liquidity measures (spread measures, share turnover, components of price impacts, and a return/volume measure). They use 18 years of intraday data for NYSE-traded companies from the ISSM Transactions database and the TAQ database, and estimate monthly time-series for the liquidity measures. Korajczyk and Sadka (2007) apply an asymptotic principal components method and find supportive evidence of common liquidity factors, which is strongest for spreads and the (fixed) components of price impacts. The first three components they compute count for more than 50 percent of the variation in individual stocks' spreads.

In contrast to the investigation of liquidity commonality across stocks, Beltran-Lopez et al. (2006) investigate liquidity commonalities in price-depth pairs in stock-specific, reconstructed Xetra order books. They apply PCA and measure liquidity by (hypothetical) price-impacts during the first three months of 2004 and confirm the existence of commonality in liquidity based on Xetra order book data. Beltran-Lopez et al. (2006) show that the first two principal components for price-impacts explain even more than 94 percent of the total variation on the individual stock level. Beltran-Lopez et al. (2006) perform PCAs on price-impacts on the bid- and ask-side of the order book separately, and find out that the two sides of the market are driven by different latent factors that drive price-impacts. Both Kempf and Mayston (2006) and Brockman and Chung (2002) investigate order-driven markets based on a similar methodology like Chordia et al. (2000), i.e. a market model for liquidity. Also Kempf and Mayston (2006) detect very strong evidence for commonality in liquidity in some stocks of the Xetra limit order market based on the PCA methodology. They stress that one shortcoming of existing studies on liquidity commonality is that only

[4]Chordia et al. (2000) argue that the inventory explanation for liquidity suggests that more trading leads to smaller spreads, since inventory risks per trade can be maintained at lower levels.

the liquidity at best bid and best ask is considered. Kempf and Mayston (2006) show that even in highly liquid order books, 20% of all orders walk up the book. Hence, there is also valuable information about liquidity beyond the first stage, in particular if there is a need to trade large volumes, where the associated orders necessarily have to walk up the order book.

2.2 Consequences of Systematic Cross-Sectional Liquidity Co-Movement

The existence of market-wide liquidity co-movements has far-reaching consequences for traders, stock exchanges, regulators, and financial theory. If the liquidity of individual stocks is at least partly determined by common (market-wide) factors, shocks to these liquidity factors influence the entire market. For the individual market participant the existence of market liquidity commonality implies that some part of liquidity risk is systematic and, hence, cannot be diversified. Market-wide liquidity (risk) becomes a priced factor. But such a market-wide liquidity (risk) factor is usually disregarded in conventional asset pricing models. Consequently, traditional diversification strategies that do not consider the second moment of liquidity, fail to diversify liquidity risk effectively (see Domowitz and Wang (2002)). Future asset pricing models will have to take into account the cross-sectional dynamics of liquidity. Since commonality shocks have consequences for the liquidity of all stocks in the market, the understanding of these market-wide liquidity determinants is important for the functioning and stabilizing of financial markets in general. Therefore, market-wide liquidity (risk) is also a policy- and regulation-issue.

3 Market Structure and Data

The *Deutsche Börse AG* runs the *Frankfurt Stock Exchange (FWB)*, which is the definitely the most important among the eight German stock exchanges, and offers floor trading as well as fully-electronic trading on the *Xetra* system. While many of the DAX-30 stocks are also listed on other exchanges, 97% of all trading of these stocks takes place in the Xetra system (see Deutsche Börse AG (2005)). Xetra operates as an open limit order book. Compared to the Paris stock exchange, not only the five best orders are displayed to the the traders. In contrast to the NYSE, every trader has insight into the complete order book. Hence, the Xetra order book is an ideal laboratory to investigate common liquidity determinants in the entire Xetra order book beyond the best-quote perspective of liquidity .

The order data-set we apply is not limited to best quotes. It consists of each single Xetra order book entry during the period 6-13 Dec 2006.[5] Each entry comprises time stamp, order number, order limit, order volume, trader and order restrictions, an indication of buyer or seller initiation, type of order entry (insertion, cancellation, execution, partial execution, system insertion/deletion), and an indication of the trading phase.

For our investigation, we aggregate the data that was originally time-stamped in 1/100-second intervals to one-minute intervals. Since the auction phases at 09:00 o'clock and 13:00 o'clock follow an entirely different trading mechanism than continuous trading, we exclude

[5] All computations are also conducted *out of sample* during the period 10-17 Jan 2007. Since the results for the two time-periods are highly consistent, we only report the results for the period 6-13 Dec 2006.

the time around auctions from the sample. The intraday time-periods investigated in this study are:

- Morning period: 10:00-11:59
- Noon period: 12:00- 12:50 and 13:10-14:59
- Afternoon period: 15:00-16:59
- Total: 10:00-12:50, and 13:10-16:59

Finally, we end up with a data-set consisting of 2394 observations for each liquidity measure during 6-13 Dec 2006. Despite the high computational effort, we run all calculations for the entire DAX-30 market and do not exclude any stocks from the sample. If systematic market-wide liquidity is investigated, the cross-section of all stocks in a market should be studied in co-variation analysis.

4 Liquidity Measurement and Extraction of Systematic Factors

4.1 Intraday Liquidity Measurement from Order Book Data

At each time t the order book for a stock i consists of the set of available ask volume A_t and the set of available bid volume B_t given by

$$A_t = \{(l^a_{tj}, v^a_{tj}) | j = 1, 2, 3, ..., n_a; l^a_{t1} < l^a_{t2} < ... < l^a_{tn^a}\} \quad (1)$$
$$B_t = \{(l^b_{tj}, v^b_{tj}) | j = 1, 2, 3, ..., n_b; l^b_{t1} > l^b_{t2} > ... > l^b_{tn^b}\} \quad (2)$$

where the stock index i is omitted. l^a_{tj} (l^b_{tj}) is the j-th best ask (bid) limit in the order book and v^a_{tj} (v^b_{tj}) the total available volume at this ask (bid) limit at time t. $l^a_{tn_a}$ ($l^b_{tn_b}$) is the highest (lowest) ask (bid) limit in the order book at time t at which volume is provided.

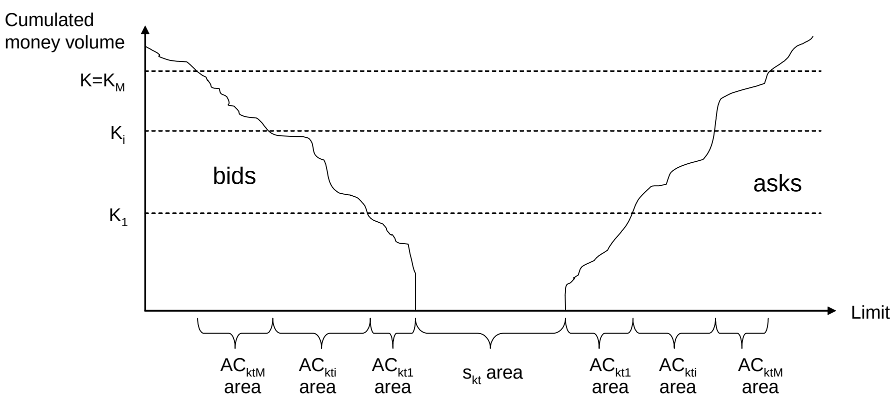

Figure 1: Order Book Illustration and Cost Attribution

We measure liquidity in terms of round-trip costs, i.e. the costs of buying K Euros of the stock at time t and selling the stocks immediately after the purchase also at time t. Like

Irvine et al. (2000), we refer to the *immediate supply of liquidity*, i.e. the cost of an immediate round-trip of the volume K Euro against the order book at time t.[6] The costs of a round-trip $RTC_{kt}(K)$ of trading K Euro at time t are given by

$$RTC_{kt}(K) = \frac{K}{\sum_{i=1}^{J^a} v_{ti}^a} - \frac{K}{\sum_{i=1}^{J^b} v_{ti}^b}, \tag{3}$$

where J^a and J^b fulfill

$$\sum_{i=1}^{J^a} v_{ti}^a l_{ti}^a = \sum_{i=1}^{J^b} v_{ti}^b l_{ti}^b = K. \tag{4}$$

For our investigation we attribute the round-trip costs to *different areas of the order book* as shown in Figure 1, which illustrates the roundtrip cost attribution for stock k at time t: spread s_{kt}, additional roundtrip costs AC_{kti} for trading a money volume of $K_i - K_{i-1}$ additional to already traded K_{i-1} Euros. The attributed part of the roundtrip costs $RTC_t(K)$ which is independent of the targeted trading volume K at a certain time t is given by the bid-ask spread s_t, i.e. by

$$s_{kt} = l_{t1}^a - l_{t1}^b. \tag{5}$$

The remaining roundtrip costs dependent on K given by

$$RTC_{kt}(K) - s_t \tag{6}$$

are attributed by considering different trading volumes $K_1 < ... < K_i < ... < K_M = K$. We calculate the ***additional roundtrip costs*** AC_{kti} of trading $K_i - K_{i-1}$ Euros additional to the already traded K_i Euros:

$$AC_{kti} = \begin{cases} RTC_{kt}(K_1) - s_t & i = 1 \\ RTC_{kt}(K_i) - RTC_t(K_{i-1}) & i > 1 \end{cases} \tag{7}$$

For the *additional round-trip cost measures* we consider several different volumes ranging from K = 25.000 to K = 1.000.000 Euro. In this paper we study the following five *non-overlapping* liquidity measures:

Table 1: Liquidity Measure Notation

Liquidity Measure	Money-Volume [Euro]	Symbol
Bid-Ask Spread	K = s	S
Additional Roundtrip Costs	K = 25.000	AC 25T
Additional Roundtrip Costs	K = 100.000	AC 100T
Additional Roundtrip Costs	K = 500.000	AC 500T
Additional Roundtrip Costs	K = 1.000.000	AC 1000T

[6] Cost of roundtrip measures were also proposed by Irvine et al. (2000), Gomber et al. (2004), and Beltran-Lopez et al. (2006).

4.2 Robust Principal Components Analysis

The PCA methodology has become the predominant method to measure latent determinants in studies on commonality in liquidity, and is also suited for an intraday investigation of the systematic part of individual stock liquidity. This method allows to judge how much systematic liquidity the applied measures show. When we run a PCA, we extract linear combinations of all individual variables according to the total variability they explain. These linear combinations are contained in the eigenvectors of the covariance matrix. But it is well documented that the results of conventional principal components analysis can be strongly influenced by the presence of outliers in the data sample (see for instance Filzmoser and Fritz (2007), and Croux and Ruiz-Gazen (2005)). More robust PCA methods arise from the Projection-Pursuit (PP) principle, which was first proposed by Huber (1985). In a nutshell, PP methods find structures in multivariate data by *projecting* the original high-dimensional data on a lower-dimensional subspace. The data is projected on a lower-dimensional space such that a *robust measure of variance of the projected data will be maximized*.

Following Croux et al. (2007), also principal components analysis is a PP method:[7] Consider n observations, and all of these observations are column vectors of dimension p. The variance is denoted by S^2. The first principal components can be extracted by finding the unit vector a that maximizes the variance of the data that gets projected. The principal components are computed by maximizing

$$a_1 = arg\ max\ S^2\,(a^t x_1, ..., a^t x_n). \tag{8}$$

From this step we obtain univariate data $a^t x_1, ..., a^t x_n$ by projecting the (multivariate) data in the direction a_1. When we apply the variance of the sample as projection index, a_1 is the eigenvector of the sample covariance matrix of the data corresponding to the largest eigenvalue (see Croux et al. (2007)).

But if we not use the sample variance as projection index, it is much more difficult to solve equation (8) and *approximate algorithms* are needed to compute the principal components. From different scale measures (like the median absolute distance) as projection index, different types of PCAs can be constructed - and much more robust results can be achieved.[8] However, our multi-stage PCA approach uses the Projection-Pursuit technique and the algorithm of Croux and Ruiz-Gazen (2005) to compute the principal components by maximizing equation (8). The structure of this algorithm is as follows: Consider a data matrix X with n rows (observations) and p columns (variables), c denotes the step of PCA. The algorithm

$$\mu_c = S^2(a_c^t x_1, ..., a_c^t x_n) \tag{9}$$

gives a sequence of approximations for the unit vectors defined in equation (8). If we use the *sample variance* for S^2, then the values of μ_c are the *eigenvalues of the covariance matrix*, ranked from the highest to the lowest (see Croux et al. (2007)). The algorithm we apply works best in cases where the sample size is relatively large compared to the number of variables, since the trial directions the algorithm considers are pointing in the directions where the data is. Unlike in conventional PCA, the data is centered with the median to get more robust results. For an exhaustive description of the algorithm see Croux and Ruiz-Gazen (2005).

[7] The notation is leaned on Croux et al. (2007).

[8] See Croux et al. (2007), pp. 2-5.

In our case the data matrix X comprises the time-series of a particular liquidity measure for all thirty stocks in the sample. If we run a PCA on this matrix, we obtain a market-wide liquidity vector M that is measure-specific. We repeat this procedure for each measure to get the corresponding vectors $M_{1...n}$ for n different measures. To extract *measure-independent* systematic liquidity factors from a second step, we again rely on Projection-Pursuit PCA for the estimation of measure-independent systematic factors. Therefore, we first isolate the time-series of the discriminant scores of the different measures' first three principal components. A matrix containing the time-series of the first three systematic factors of each of the five liquidity measures applied is the input into the second PCA stage. We can now differentiate systematic liquidity factors that are not only non-idiosyncratic, but also independent of the underlying non-overlapping liquidity measures.

5 Measure-Specific Systematic Liquidity Factors

5.1 Cross-Sectional PCA on Non-Overlapping Liquidity Measures

We denote the mean adjusted spread of stock i at time t by s_{it} and the mean adjusted additional trading costs for this stock by AC_{mit}. As we consider the DAX-30 stock market, $i = 1, ..., 30$. For investigating commonality in liquidity we aggregate n stock liquidity measures in a first step over all stocks such that we have $M + 1$ corresponding K-dimensional market wide liquidity measure vectors for each t which k-th component is given by

$$s_t^k = \sum_{i=1}^{30} \lambda_{ik}^s s_{it} \, , \; and \tag{10}$$

$$AC_{1t}^k = \sum_{i=1}^{30} \lambda_{ik}^{AC_1} AC_{1it} \tag{11}$$

$$\cdots$$

$$AC_{Mt}^k = \sum_{i=1}^{30} \lambda_{ik}^{AC_M} AC_{Mit}, \tag{12}$$

where the λ_{ik}s denote the corresponding discriminant scores of component k for stock i.

Table 2 gives the results of the intraday cross-sectional principal components analysis for the five liquidity measures applied.[9] The table shows the proportions of explained and cumulative explained variance for the first three systematic factors of each liquidity measure in the time period 10:00-12:50 and 13:30-16:59 during 6-13 Dec 2006. The results in the table illustrate that the liquidity of individual stocks in the market is considerably driven by latent market-wide factors. Regardless which measure we apply, the first three systematic factors explain more than 40% of the (measure-specific) liquidity variation in individual stocks.

Since most researchers that apply PCA on liquidity-related measures rely on bid-ask spread measures, we discuss our PCA results for the bid-ask spread in more detail. The first three common factors resulting from our PCA explain more than 46% of the total variation

[9] The corresponding factor loadings of the stocks' liquidity measures towards their systematic liquidity factors can be found in the appendix section.

Table 2: Explained Variances of Systematic Liquidity Factors

Measure		M [1]	M [2]	M [3]
S	Proportion of Variance	16.25 %	15.36 %	14.94 %
	Cum.	16.25 %	31.61 %	46.55 %
AC 25T	Proportion of Variance	15.17 %	14.00 %	12.40 %
	Cum.	15.17 %	29.17 %	41.56 %
AC 100T	Proportion of Variance	21.01 %	20.04 %	13.31 %
	Cum.	21.01 %	41.05 %	54.36 %
AC 500T	Proportion of Variance	21.52 %	20.79 %	11.66 %
	Cum.	21.52 %	42.31 %	53.97 %
AC 1000T	Proportion of Variance	20.99 %	16.47 %	10.61 %
	Cum.	20.99 %	37.46 %	48.08 %

The summary table shows the proportions of explained and cumulative explained variance of individual stocks' liquidity for the first three principal components of five liquidity measures during the period 6-13 Dec 2006. The calculation is based on projection-pursuit principal components analysis (PP-PCA) using the Croux/Ruiz (2005) algorithm.

in individual stocks' spreads. Hasbrouck and Seppi (2001) find an explanatory power above 50% of the total variation for the first three components. Also Korajczyk and Sadka (2007) detect supportive evidence of common liquidity factors in the case of spread measures. Also in their study, the first three components count for more than 50% of the variation in the spreads of individual stocks. However, the low explanatory power of the first component of the bid-ask spread (relative to the other systematic factors) is also documented in the study of Kempf and Mayston (2006).

Concerning the roundtrip cost measures AC, we detect even stronger evidence of commonality. The commonality exhibited in roundtrips of $K = 25T \ldots 1000T$ Euro is (except for AC 25K) considerably higher than for best-quote liquidity commonality described by the bid-ask spread. Also note that the commonality in the measures AC 100T and AC 500T is highly comparable, and higher than for volumes $K < 100.000$ and $K > 1.000.000$ Euro.

These results also hold for different times during the trading day. Detailed results for *different time periods during the trading session* (morning, noon, and afternoon) can be found in the appendix section. From the investigation of morning, noon, and afternoon periods, we can conclude that commonality varies during the day. In more detail, there is a variation in the explained variances of the single latent factors, but the cumulated proportion of explained variance for the first three systematic liquidity factors is quite constant during the different time periods. One exception is the commonality in the measure AC 100T during the morning hours, where the cumulative proportion the first three systematic factors explain is nearly 10% higher than during the noon or afternoon hours. AC 100T and AC 500T show the highest commonality across stocks, with a maximum of around 60% cumulative explained variance of the measure AC 500T during the morning time period. Hence, we can conclude from the investigation of different time periods during continuous trading that there are fluctuations in the variance the different systematic factors explain. But, overall, the cumulated proportions remain quite stable throughout the trading day.

5.2 Canonical Correlations between Systematic Liquidity Factors

From PCA we have already extracted latent market-wide determinants of the liquidity of individual stocks in the market. One question that arises is whether the measure-specific systematic factors extracted by the use of PCA are related to each other, i.e. do the systematic liquidity factors show correlation in the cross-section. Therefore, we conduct canonical correlation analysis on the different systematic liquidity factors to get a more detailed picture of interactions across different liquidity measures.

For instance, canonical correlation analysis is also applied by Hasbrouck and Seppi (2001) and Korajczyk and Sadka (2007). While Hasbrouck and Seppi (2001) compute the canonical correlations between order imbalances and stock returns to examine if the the two variables show commonality, we study co-variation across systematic liquidity factors. Similar to Korajczyk and Sadka (2007), we investigate the systematic factors' correlations across each pair of the different liquidity measures.

Canonical correlation analysis allows the investigation of the relation between two sets of variables: Let E_t and F_t denote two sets of variables, which are in our case the discriminant scores of the principal components for each pair of liquidity measures. The first canoncial variates are the linear combinations $a\,E_t$ and $b\,F_t$ such that the correlation between the linear compounds is maximized: $max\,Corr(a\,E_t + b\,F_t)$.

Table 3: Canonical Correlations during 6-13 Dec 2006

	AC 25T	AC 100T	AC 500T	AC 1000T
S	0.29	0.35	0.38	0.21
	0.25	0.34	0.36	0.17
	0.19	0.31	0.31	0.15
AC 25T		0.57	0.31	0.29
		0.31	0.18	0.22
		0.21	0.16	0.16
AC 100T			0.43	0.29
			0.32	0.20
			0.29	0.18
AC 500T				0.38
				0.35
				0.26

The summary table reports the first three canonical correlations between each pair of the underlying systematic liquidity factors during the period 6-13 Dec 2006.

Table 3 gives the first three canonical correlations between each pair of liquidity measures. The most obvious, and also the most interesting result is that changes in the systematic factors of the different liquidity measures are correlated. This means that also the systematic components of different measures are correlated across different measures of liquidity. The highest canonical correlations can be found between the different AC measures.

The weakest correlations are between the discriminant scores of the bid-ask spread and AC 1000T, probably due to the fact that the two variables measure liquidity in areas, which are far away from each other. One pattern that can be identified is that each AC measure is in any case most strongly correlated with the measure with the next higher or lower target money-volume K. Interestingly, the canonical correlation between systematic factors in the bid-ask spread, AC 100T, and AC 500T are considerably higher than for K=25T, although the order book areas the measures capture are further away from the bid-ask spread than AC 25T. The finding that the changes in the systematic part of different liquidity measures are correlated across measures are in a line with the results of Acharya and Pedersen (2005) and Korajczyk and Sadka (2007).

5.3 Canonical Correlations with other Trading-Related Measures

In this section, we investigate whether the latent liquidity determinants already extracted are also related to the systematic components of other trading-related variables. First, we investigate a volume-based measure that informs about the maximum money-volume K_{it}^{max} of stock i that is *immediately tradable* at time t, where the associated order does not eat into the order book, i.e. the cost of the trade is only the bid-ask spread. This trading-related measure is defined as the aggregated money-volume of all shares quoted at best bid and best ask at time t

$$VS_{it}^{max} = n_{it}^{b} * l_{it}^{b} + n_{it}^{a} * l_{it}^{a}, \tag{13}$$

where n_{it}^{b} (n_{it}^{a}) is the number of shares quoted at best bid (ask), and l_{it}^{b} (l_{it}^{a}) denotes the best bid (ask) limit for stock i at time t.

As a return/volume measure we apply a variable that was originally inspired by the *Illiquidity Measure* proposed by Amihud (2002). He measures daily illiquidity as the "... average ratio of daily absolute return on the (dollar) trading volume on that day (...)."[10]. He argues that this ratio gives the "daily price impact of the order flow."[11] Conversely, we apply an *intraday price/volume measure* PV_{it} in one-minute intervals. This ratio for stock i at time t is given by the equation

$$PV_{it} = \frac{l_{it}^{mid}}{l_{it}^{mean.a} * A_{it} + l_{it}^{mean.b} * B_{it}} \tag{14}$$

$$= \frac{0.5 * (l_{it}^{a} + l_{it}^{b})}{l_{it}^{mean.a} \sum_{i=1}^{N_t^b} v_{it}^{b} + l_{it}^{mean.a} \sum_{i=1}^{N_t^a} v_{it}^{a}}, \tag{15}$$

where $l_{it}^{mean.a}$ ($l_{it}^{mean.b}$) is the average limit, i.e. the arithmetic mean, of all quoted limits on the ask-side (bid-side) of the order book, and l_{it}^{a} (l_{it}^{b}) denotes the best ask limit (bid limit) for stock i at time t. The mid-limit l_{it}^{mid} is the mid-point between best bid and best ask for stock i at time t:

$$l_{it}^{mid} = l_{it}^{a} - l_{it}^{b})/2\,. \tag{16}$$

[10] Amihud (2002), p. 34.

[11] Amihud (2002), p. 34. Hasbrouck (2005) notes that Amidud's *Illiquidity Measure* works better than other measures in capturing the price impact parameter λ proposed by Kyle (1985), i.e. the response of price to order-flow.

Since we compute all measures dynamically, i.e. by changes, our intraday price/volume measure PV_{it} captures the percentage change of the mid-limit in relation to the quoted volumes weighted by the mean-limit on both sides of the order book in one-minute intervals.

The next trading-related measure we investigate is relative *volume imbalance* IM_{it}, which we define as the relative difference between the total ask-side volume and the total bid-side volume of the order book at time t. Chordia and Subrahmanyam (2004) argue that "(...) there are at least two reasons why order imbalances can provide additional power beyond trading activity measures such as volume in explaining stock returns. First, a high absolute order imbalance can alter returns as market makers struggle to re-adjust their inventory. In addition, order imbalances can signal excessive investor interest in a stock (...)."[12] We also argue that imbalances between the two sides of the market may offer valuable information about stock or market sentiment. Imbalances between the two sides of the order book, or massive shifts in imbalance mean that there exists demand- or supply-pressure, which may indicate a (future) rise or fall in transaction prices. While Chordia and Subrahmanyam (2004) examine imbalances in the number and the money-volume of *transactions* and investigate the time-series relation between order imbalances and individual stock returns, we measure *imbalances in quoted volumes between the two sides of the market* and extract their systematic factors to investigate the relation between systematic imbalances and systematic liquidity. However, our measure captures imbalances between the cumulated volumes on the two sides of the entire order book, which measures excess demand or supply based on all limit orders in the book. We define (relative) order volume imbalance IM_{it} for stock i as the difference between the volumes on the demand and supply side of the order market, relative to the total order volume quoted in the order book at time t

$$IM_{it} = 2\,\frac{A_{it} - B_{it}}{A_{it} + B_{it}}, \tag{17}$$

where B_{it} (A_{it}) is the total order volume on the bid (ask) side of the order book of stock i at time t.[13] For dealing with cumulative probabilities from discrete distributions, see Paarsch and Hong (2006).

Table 4 shows the first three canoncial correlations between the systematic liquidity factors in roundtrip-based measures and the systematic factors of other trading-related measures.[14] The canonical correlations between the money-volume quoted at the spread, i.e. the maximum money-volume that is immediately tradable where the price of the trade is only the bid-ask spread, and the changes in the bid-ask spread are only weak. Overall, the money-volume quoted at the bid-ask spread does not correlate considerably with the AC liquidity measures. Conversely, we detect a *strong relation between systematic factors in our price/volume measure and the systematic factors of all roundtrip-based measures* applied. Not surprisingly, there is no strong relation between the systematic factors of changes in volume imbalances and the market-wide factors of the bid-ask spreads, since the two measures cover entirely different order book areas. The strong relation between systematic factors in order imbalances between the two sides of the market and the market-wide determinants of additional round-trip costs is one of the most interesting findings of this section: This

[12] Chordia and Subrahmanyam (2004), p. 486.

[13] The ratio is multiplied by the factor 2 due to reasons of numerical computation.

[14] Note that we extract the systematic factors of these trading-related variables by the use of PP-PCA that, again, relies on the Croux/Ruiz (2005) algorithm.

empirical fact means that *changes in market-wide demand or supply pressure are strongly related to changes in the systematic part of individual stocks' liquidity*. The correlations are considerably higher in the anterior regions of the order book, i.e. for K $\leq$ 100.000 Euro, where most orders activities takes place. Concerning systematic changes in mid-limits we find out that there is only a weak correlation with the systematic factors exhibited in the most popular liquidity measure bid-ask spread. But we detect very strong correlations between the market-wide determinants of individual stocks' mid-limit changes and systematic changes in additional round-trip cost measures. Consequently we can conclude that *the systematic fraction of individual stocks' mid-limits is strongly related to changes in systematic liquidity*, and that this relation is even more severe with systematic liquidity in the anterior regions of the order book.

Table 4: Canonical Correlations between Systematic Factors

Measure	Bid-Ask Spread	AC 25T	AC 100T	AC 500T	AC 1000T
Money-Volume at Spread	0.15	0.15	0.15	0.13	0.14
	0.13	0.12	0.12	0.12	0.12
	0.12	0.10	0.11	0.11	0.10
Price/Volume	0.18	0.61	0.48	0.55	0.53
	0.15	0.54	0.34	0.51	0.46
	0.13	0.34	0.27	0.47	0.44
Volume Imbalance	0.13	0.41	0.45	0.26	0.25
	0.11	0.38	0.17	0.17	0.22
	0.09	0.30	0.12	0.13	0.17
Mid-Limit	0.16	0.90	0.70	0.41	0.32
	0.14	0.63	0.46	0.23	0.29
	0.12	0.54	0.15	0.14	0.20

The summary table reports the first three canonical correlations between the systematic liquidity components of five liquidity-measures and four trading-related variables during the period 6-13 Dec 2006.

We can summarize section 5 with the following findings: (1) Based on highly robust cross-sectional principal components analysis of five non-overlapping liquidity measures, we detect heavy-weighting market-wide liquidity determinants. Regardless which liquidity measure we apply, the first two or three principal components explain more than 50% of the variation in individual stocks' liquidity in the market. So, the single-asset perspective on liquidity is not able to effectively examine liquidity (risk) and, therefore, no longer sustainable. Future asset pricing models need to consider cross-sectional liquidity interactions in a market. (2) From canonical correlations analysis we show that also the systematic liquidity components are correlated across different non-overlapping liquidity measures. (3) Concerning the additionally applied trading-related measures, we find out that the systematic liquidity factors are highly correlated with changes in market-wide demand or supply pressure. Moreover, we detect strong correlations between the market-wide determinants of individual stocks' mid-limit changes and systematic changes in liquidity in the anterior areas of the order book.

6 Global Systematic Liquidity

6.1 Aggregation of Measure-Specific Systematic Liquidity Factors

Most researchers study commonality with a single measure or with different measures *separately*. Such approaches only allow to detect common liquidity factors that are dependent on the particular facet of liquidity the applied measure is able to capture. Well, some try to examine market-wide liquidity (risk) factors over several different liquidity variables (p.e. Korajczyk and Sadka (2007)), which measure mixed liquidity-related aspects in a more or less undifferentiated manner. An estimation of systematic liquidity from overlapping measures is problematic and may give deceiving results that lead to misinterpretations. To overcome these methodological drawbacks, we examine whether there are also market-wide liquidity determinants extractable, which are not only stock-independent but also *independent of the underlying **non-overlapping** liquidity measures*. In the following we call these latent determinants *global systematic liquidity factors*.

6.2 Global Systematic Liquidity Factors

We aggregate the five measure-specific market-wide liquidity measure vectors $M_{1...5}$ to a multi-dimensional *global* market-wide liquidity vector G. To extract *measure-independent* systematic liquidity factors, we again rely on on the same PP-PCA approach for the estimation of the latent factors. First, we isolate the time-series of the discriminant scores of the different measures' first three principal components obtained from the first cross-sectional PCA.[15] The input into the second PCA is a matrix containing the discriminant scores' time-series of the first three systematic factors of each of the five non-overlapping liquidity measures applied. This step gives the systematic liquidity factors $G_{1...n}$, which are (1) stock-independent, and (2) also independent of the underlying measure.

Therefore, we aggregate the market-wide liquidity measure vectors $M_{1,...,n}$ to a H-dimensional *global* market-wide liquidity vector G_t, which h-th component is given by

$$\begin{aligned} G_t^h &= \sum_{k=1}^{K} \xi_{hk}^s s_t^k + \sum_{k=1}^{K} \xi_{hk}^{AC_1} AC_{1t}^k + ... + \sum_{k=1}^{K} \xi_{hk}^{AC_M} AC_{Mt}^k \\ &= \sum_{k=1}^{K} \left(\xi_{hk}^s \sum_{i=1}^{30} \lambda_{ik}^s s_{it} + \xi_{hk}^{AC_1} \sum_{i=1}^{30} \lambda_{ik}^{AC_1} AC_{1it} + ... + \xi_{hk}^{AC_M} \sum_{i=1}^{30} \lambda_{ik}^{AC_M} AC_{Mit} \right) \\ &= \sum_{i=1}^{30} \left(s_{it} \sum_{k=1}^{K} \xi_{hk}^s \lambda_{ik}^s + AC_{1it} \sum_{k=1}^{K} \xi_{hk}^{AC_1} \lambda_{ik}^{AC_1} + ... + \sum_{k=1}^{K} \xi_{hk}^{AC_M} \sum_{i=1}^{30} AC_{Mit} \lambda_{ik}^{AC_M} \right) \\ &= \sum_{i=1}^{30} \left(s_{it} \gamma_{hi}^s + AC_{1it} \gamma_{hi}^{AC_1} + ... + AC_{Mit} \gamma_{hi}^{AC_M} \right), \end{aligned} \tag{18}$$

where the ξs denote the corresponding factors loadings of the market-wide liquidity vectors $M_{1,...,n}$.

[15] In this study, we concentrate on the three most important systematic measure-independent liquidity factors G_1, G_2, and G_3, since these three factors show a cumulative proportion of explained variance of more than 70% of the liquidity variation in individual stocks.

The results in table 5 are based on Projection-Pursuit PCA on the first three systematic (measure-dependent) factors of each liquidity measure. Explained and cumulative explained variance as well as the systematic factors' standard deviations are reported in the table for different time-periods during the trading session. The *morning period* is during 10:00-11:59, the *noon period* comprises the periods 12:00-12:50 and 13:10-14:59, the *afternoon period* is during 15:00-16:59, and the *total* time-period comprises the periods 10:00-12:50 and 13:10-16:59. Our most important result is that we can demonstrate that there are systematic liquidity factors, which are also independent of underlying non-overlapping liquidity measures. The first three global systematic liquidity factors we extract explain around 70% of the total variation in the first three measure-specific common liquidity factors of each measure.

Table 5: Explained Variances of Global Liquidity Factors

Time Period		G [1]	G [2]	G [3]
Morning Period	Standard Deviation	8.82	7.14	6.04
	Proportion of Variance	**33.90** %	22.21 %	15.91 %
	Cum.	33.90 %	56.11 %	72.03 %
Noon Period	Standard Deviation	6.09	5.65	5.99
	Proportion of Variance	27.07 %	23.29 %	18.21 %
	Cum.	27.07 %	50.27 %	68.58 %
Afternoon Period	Standard Deviation	8.98	8.00	6.73
	Proportion of Variance	29.60 %	23.50 %	16.61 %
	Cum.	29.60 %	53.10 %	69.72 %
Total	Standard Deviation	7.31	6.68	5.86
	Proportion of Variance	28.33 %	23.68 %	18.18 %
	Cum.	28.33 %	52.00 %	70.18 %

The summary table shows the proportions of explained and cumulative explained variance for the global principal components during the period 6-13 Dec 2006. The calculation is based on projection-pursuit principal components analysis using the Croux/Ruiz (2005) algorithm.

Concerning the *different time periods* during the trading session, the picture is comparable to the results obtained from the first PCA stage. While the proportions of variance the different systematic factors explain change during the day, the cumulative proportion of variance the first three global systematic factors explain is very stable, and varies only between 68.58% during the noon period and 72.03% during the morning hours. A finding that longs for a more detailed investigation is that the explanatory power of G [1] falls considerably from 33.90% during the morning period to only 27.07% during the noon time period, while the other two factors stay relatively stable throughout the day.[16] We can support this finding with the same observation in a 'real' out of sample test during the period 10-17 Jan 2007.

Table 6 reports the factor loadings of each of the first three principal components of measure-specific systematic liquidity factors towards the three global systematic liquidity factors. Concerning the *first global liquidity factor* G [1], the market-wide factors S [1], AC

[16]What determines the strong variation of the first global systematic liquidity factor G [1] is investigated in section 6.3.

500T [1], and AC 500T [2] show the strongest correlations with G [1] (> 0.40). Weaker, but also considerable loadings towards G [1] can be found in the factors AC 100T [2], AC 500T [3], AC 1000T [1], and AC 1000T [2]. Concerning the *second global factor* G [2] the strongest relations can be found in S [1], AC 500T [1], and AC 500T [3], which are all above 0.40. Weaker factors loadings with G [2] are exhibited in S [2], S [3], and AC 500T [1]. Note that the loadings of the systematic liquidity factors AC 25T and AC 1000T show only very little loadings towards the second global liquidity factor G [2]. In the case of the *third global global liquidity factor* G [3] the picture changes. The highest loadings can be found in the systematic factors of S, AC 500T, and AC 1000T, whereas AC 25T and AC 100T show very weak loadings towards G [3].

Table 6: Loadings towards Global Factors

Liquidity Measure		G [1]	G [2]	G [3]
S	M [1]	0.402	-0.425	-0.244
	M [2]	0.028	-0.111	0.420
	M [3]	-0.034	-0.263	0.094
AC 25T	M [1]	-0.037	0.011	-0.007
	M [2]	0.001	-0.006	-0.010
	M [3]	0.019	-0.008	0.012
AC 100T	M [1]	0.003	-0.144	0.010
	M [2]	0.119	-0.036	0.041
	M [3]	0.019	-0.009	0.042
AC 500T	M [1]	-0.531	-0.654	0.222
	M [2]	0.670	-0.089	0.555
	M [3]	-0.189	0.515	0.200
AC 1000T	M [1]	-0.203	0.088	0.547
	M [2]	0.107	-0.022	-0.201
	M [3]	-0.003	-0.087	-0.118

Factor loadings of the first three systematic liquidity factors of the five liquidity measures and the global systematic liquidity factors during the period 6-13 Dec 2006.

6.3 Systematic Liquidity Factors and Global Systematic Liquidity Factors

To get a more detailed picture of the relations between sytematic, measure-specific liquidity factors and the global liquidity factors, we next conduct uni- and multivariate regressions between the measure-dependent, market-wide liquidity factors (resulting from first-stage PCA) and the global systematic liquidity measures (resulting from the second PCA). Since the first three systematic factors of the different liquidity measures explain in about the half of the total liquidity variation in individual stocks, we again concentrate on these most influential factors. In more detail, we regress the systematic market-wide liquidity factors exhibited in the different liquidity measures on each of the three global liquidity factors extracted by the second cross-sectional PCA.

The corresponding equation for the regression on the first global systematic liquidity factor $G\,[1]$ is specified by

$$
\begin{aligned}
G\,[1] \quad = \quad & \beta_{Intercept} + & (19)\\
& \beta_{S\,[1]}\,S[1] + \beta_{S\,[2]}\,S[2] + \beta_{S\,[3]}\,S[3] + & (20)\\
& \beta_{AC\,25T\,[1]}\,AC\,25T[1] + \beta_{AC\,25T\,[2]}\,AC\,25T[2] + \beta_{AC\,25T\,[3]}\,AC\,25T[3] + & (21)\\
& \beta_{AC\,100T\,[1]}\,AC\,100T[1] + \beta_{AC\,100T\,[2]}\,AC\,100T[2] + \beta_{AC\,100T\,[3]}\,AC\,100T[3] + & (22)\\
& \beta_{AC\,500T\,[1]}\,AC\,500T[1] + \beta_{AC\,500T\,[2]}\,AC\,500T[2] + \beta_{AC\,500T\,[3]}\,AC\,500T[3] + & (23)\\
& \beta_{AC\,1000T\,[1]}\,AC\,1000T[1] + \beta_{AC\,1000T\,[2]}\,AC\,1000T[2] + \beta_{AC\,1000T[3]}\,AC\,1000T[3] + \epsilon, & (24)
\end{aligned}
$$

where the βs denote the corresponding multivariate regression coefficients (Coef.m), and ϵ is the error term. The regression equations for $G\,[2]$ and $G\,[3]$ are computed analogously. The models are tested for non-normality, for misspecification error, for non-constant error variance and for multicollinearity.

Table 7: Regression Statistics of Market-Wide Liquidity Factors on the First Global Liquidity Factor

		Coef.m	Sig.m	Coef.u	Sig.u	R2.u
S	M [1]	0.48	***	0.80	***	0.00
	M [2]	0.38	***	-0.13	**	0.00
	M [3]	-0.46	***	-0.29	***	0.00
AC 25T	M [1]	-0.49	***	-0.13	.	0.00
	M [2]	0.18	***	-0.03	-	0.00
	M [3]	0.25	***	-0.14	.	0.00
AC 100T	M [1]	0.13	***	-0.69	***	0.07
	M [2]	0.43	***	1.01	***	0.14
	M [3]	0.25	***	0.33	**	0.00
AC 500T	M [1]	-1.95	***	-0.80	***	0.56
	M [2]	2.46	***	1.04	***	0.69
	M [3]	-0.69	***	-0.27	***	0.02
AC 1000T	M [1]	-0.51	***	-0.51	***	0.12
	M [2]	0.39	***	0.53	***	0.06
	M [3]	-0.02	***	-0.02	-	0.00
Intercept		0.92	***			

Regression of the first three systematic factors of each liquidity measure on the *first global liquidity factor* during 6-13 Dec 2006.

Summary table 7 reports multivariate and univariate regression results. The left side of the table comprises the regression coefficient (Coef.m) and the corresponding significance code (Sig.m) in multivariate regression.[17] The columns on the right side of the table show the results of univariate regression analysis between the single market-wide liquidity components and the global liquidity factor G [1]. Reported are the regression coefficient (Coef.u), its significance (Sig.u), and the associated adjusted R^2 (R2.u) in univariate regression.

As already mentioned in section 5, we are interested in the strong *variation of G [1] during the trading session* and investigate its determinants by univariate and multivariate regression

[17] The significance codes follow the denotation: 0.001 '***'; 0.01 '**'; 0.05 '*'; 0.1 '.'; 1 '-'.

analysis using the market-wide liquidity components M of the five liquidity measures as regressors.

Since the *multivariate* part corresponds to an 'intra-model' regression, all multivariate regression coefficients are highly significant.[18] The first two systematic measure-dependent liquidity components M [1] and M [2] exhibited in the bid-ask spread (S) show a positive relation to the first global liquidity factor, whereas M [3] is negatively related to G [1]. Concerning the factors dominating the anterior regions of the order book, we can summarize that only the first measure-specific market-wide liquidity factor of AC 25K is negatively related to G [1], whereas all other systematic components are positively related to G [1]. Concerning the systematic components in AC measures in the deeper areas of the order book, i.e. for volumes $>$ 500.000 Euro, we detect the strongest relations with G [1] in AC 500T [1-3] and AC 1000T [1].

From *univariate* analysis of the single regressors, we get a clear picture of which factors dominate (the variation of) G [1]. The first two market-wide, measure-specific systematic liquidity componets exhibited in price-impacts of order-volumes of 500.000 Euro show by far the strongest explanatory power measured by the adjusted R^2: 0.69 for M [2] of AC 500T, and 0.56 for M [1] of AC 500T. Note that both are at a significance level of 99.9%. Consequently, we can argue that these two systematic (measure-dependent) liquidity components count for the major part of the variation in G [1]. Hence the extraordinarily high variation of the first global liquidity factor during the trading session is determined by latent, market-wide liquidity components exhibited in orders of high volumes, i.e. where the targeted volume K is half a million Euro.

When we regress the same market-wide liquidity factors M of the different liquidity measures on the *second and on the third global liquidity factors* (see tables 10 and 11 in the appendix section), we do not detect such strong influences. Concerning the determinants of G [2] we find one exception in the case of the first systematic liquidity component of AC 500T where the R^2 in univariate regression is 65%. However, all other factors do not show comparable R^2s. Regarding G [3] we detect even lower explanatory power of the regressors in univariate regression (see table 11 in the appendix section).

We conclude section 6 with the findings that (1) measure-independent *global systematic liquidity factors* exist, and that the first three global liquidity factors we extract explain around 70% of the total liquidity variation on the sub-level. (2) From the investigation of different time periods during the trading session, we conclude that also the global systematic liquidity factors vary in their proportion of explained variance throughout the trading day: While the proportion of variance the single factors explain fluctuate during the day (in particular G [1]), the cumulative proportion of variance the first three global factors explain stays quite stable throughout the trading session. (3) Another interesting finding from the systematic liquidity factors' factor loadings towards the global liquidity factors is that the latent factors that drive the bid-ask spreads of individual stocks show stronger correlations than the systematic parts of the AC measures that capture liquidity in the anterior regions of the order book. For targeted volumes $>$ 500.000 Euro (AC 500T and AC 1000T) the factor loadings are considerably higher than for systematic factors associated with orders that only walk up the anterior areas of the order book (AC 50T and AC 100T). (4) From regression analysis we can show that the first global factor G [1] is primarily determined by latent market-wide factors exhibited in high-volume orders that have a target volume of 500.000 Euro.

[18] Note that all liquidity measures have the same scale (Euro).

7 Conclusion

To overcome the methodological drawbacks of existing studies on market-wide liquidity co-variation, we propose a multi-stage PCA and regression approach. Using a highly robust principal components analysis method based on the Projection-Pursuit principle, we can extract reliable common factors in liquidity. The advantage of the application of non-overlapping liquidity measures, which refer to the cost of an immediate trade of different money-volumes, is that we are able to measure commonality in liquidity in a 'cleaner', and also more plausible manner.

From several cross-sectional Projection-Pursuit PCAs on different non-overlapping liquidity measures we detect heavy-weighting latent factors that determine the changes in the liquidity (measures) of individual stocks in the market. Regardless which liquidity measure we apply, the first three systematic liquidity factors explain more than 40% of the total (measure-specific) variation in individual stocks' liquidity.

The second PCA stage allows extracting systematic liquidity factors that are not only non-idiosyncratic, but also independent of the underlying, non-overlapping measures. From this step we obtain global systematic liquidity components, where the first three of them count for more than 70% of the variation on the sub-level.

We argue that traders are most severely exposed to changes in market-wide liquidity when there is a need for trading high volumes during the trading session. Therefore, we investigate three different time periods during continuous trading, and show that there is a variation in the proportions of liquidity variance the single factors explain. The cumulative proportion of variance the first three systematic factors explain stays quite stable throughout the day. These results hold for both PCA stages, i.e. for measure-dependent systematic liquidity determinants and also for the global systematic factors. The intraday-variation is due to the first global systematic liquidity factor, which is mainly determined by changes in latent measure-specific determinants exposed in price-impacts of large orders with a target volume of half a million Euro.

The investigation of the canonical correlations between the systematic factors of different liquidity measures leads us to the conclusion that also the market-wide liquidity factors of the different non-overlapping liquidity measures are correlated in the cross-section. The correlations are strongest between variables that measure liquidity in neighboring areas of the order book. Additionally, we apply four trading-related measures and investigate their relations to the the liquidity measures. Here we detect high correlations between our intraday price/volume measure and the cost of round-trip measures. A highly intersting result is the strong correlation between the latent determinants of order-volume imbalances between demand- and supply-side of the market and the systematic factors in additional round-trip cost measures. This empirical finding means that changes in market-wide demand- or supply-pressures are strongly related to systematic changes in individual stocks' liquidity. These correlations are strongest in the anterior regions of the order book where the major part of order activities take place. As a price-related measure we also study systematic changes in mid-limits and find formidable correlations with changes in systematic factors of additional round-trip costs of up to 90%.

The latest developments in the literature on market-wide liquidity, in particular the actual results obtained from the investigation of order book data, have shown that latent market-wide factors exist, and that these factors systematically drive the liquidity of individual stocks. In most studies, more than the half of the variation of individual stocks' liquidity can be explained by the first three systematic factors. The convincing evidence of commonality in liquidity indicates that the predominant single-stock perspective on liquidity is no longer sustainable. The fact that the liquidity of individual stocks is such severely driven by latent market-wide factors longs for enhanced asset pricing models that also consider the second moment of liquidity.

References

Acharya, V. and Pedersen, L. (2005). Asset pricing and liquidity risk. *Journal of Financial Economics*, 77:275–310.

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

Beltran-Lopez, H., Giot, P., and Grammig, J. (2006). Commonalities in the order book. *CORE Working Paper, March 2006, Center for Operations Research and Econometrics*.

Brockman, P. and Chung, D. (2002). Commonality in liquidity: Evidence from an order-driven market structure. *Journal of Financial Research*, 25:521–539.

Chordia, T., Roll, R., and Subrahmanyam, A. (2000). Commonality in liquidity. *Journal of Financial Economics*, 56:3–28.

Chordia, T. and Subrahmanyam, A. (2004). Order imbalance and individual stock returns: Theory and evidence. *Journal of Financial Economics*, 72:485–518.

Croux, C., Filzmoser, P., and Oliveira, M. (2007). Algorithms for projection-pursuit robust principal component analysis, chemometrics and intelligent laboratory systems. available at: http://www.statistik.tuwien.ac.at/public/filz/papers/pppca07.pdf.

Croux, C. and Ruiz-Gazen, A. (2005). High breakdown estimators for principal components: the projection-pursuit approach revisited. *Journal of Multivariate Analysis*, 95:206–226.

Deutsche Börse AG (2005). Order_book_statistics_xetra_close-1.xls. *Online Publication*. online available at: http://deutsche-boerse.com.

Domowitz, I. and Wang, X. (2002). Liquidity, liquidity commonality and its impact on portfolio theory. *SSRN Working Paper Series, Working Paper*, 296870.

Filzmoser, P. and Fritz, H. (2007). Exploring high-dimensional data with robust principal components. *Forschungsbericht*, CS-2007-2.

Gomber, P., Schweickert, U., and Theissen, E. (2004). Zooming in on liquidity. *EFA 2004 Maastricht Meeting Papers, Paper No. 1805, also available at SSRN*.

Hasbrouck, J. (2005). Trading costs and returns for us equities: The evidence from daily data. *Stern School Working Paper, 3rd Preliminary Version.*

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

Huber, P. (1985). Projection pursuit. *The Annals of Statistics*, 06/04.

Irvine, P., Benston, G., and Kandel, E. (2000). Liquidity beyond the inside spread: Measuring and using information in the limit order book. *Available online: http://gbspapers.library.emory.edu/archive/00000159/01/GBS-FIN-2000-002.pdf*.

Kempf, A. and Mayston, D. (2006). Systematic liquidity across limit order books: Evidence from order data. *CFR Working Paper*, 06-04.

Korajczyk, R. and Sadka, R. (2007). Pricing the commonality across alternative measures of liquidity. *SSRN Working Paper Series, Working Paper*, 900363.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335.

O'Hara, M. (1995). *Market Microstructure Theory*. Blackwell Publishers, Cambridge, USA.

Paarsch, H. and Hong, H. (2006). *An Introduction to the Structural Econometrics of Auction Data*. The MIT Press.

Zheng, X. and Zhang, Z. (2006). Commonality in liquidity in emerging markets: Evidence from the chinese stock market. *University of Durham Working Papers in Economics and Finance*, 06/04.

A Appendix

Table 8: Explained variances resulting from PP-based PCA using the Croux/Ruiz (2005) algorithm during 6-13 Dec 2006

Measure		Total			Morning			Noon			Afternoon		
	Component	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]
S	Proportion of Variance	16.25 %	15.36 %	14.94 %	16.94 %	15.91 %	15.82 %	16.39 %	14.22 %	13.95 %	17.52 %	15.01 %	14.14 %
	Cum.	16.25 %	31.61 %	46.55 %	16.94 %	32.84 %	48.66 %	16.39 %	30.61 %	44.56 %	17.52 %	32.53 %	46.67 %
AC 25T	Proportion of Variance	15.17 %	14.00 %	12.4 %	16.88 %	13.12 %	13.06 %	16.16 %	15.92 %	12.93 %	15.94 %	13.98 %	13.85 %
	Cum.	15.17 %	29.17 %	41.56 %	16.88 %	30.00 %	43.06 %	16.16 %	32.08 %	45.02 %	15.94 %	29.92 %	43.77 %
AC 100T	Proportion of Variance	21.01 %	20.04 %	13.31 %	26.14 %	24.93 %	9.36 %	18.99 %	18.26 %	13.85 %	20.41 %	19.31 %	11.16 %
	Cum.	21.01 %	41.05 %	54.36 %	26.14 %	51.07 %	60.43 %	18.99 %	37.25 %	51.1 %	20.41 %	39.72 %	50.88 %
AC 500T	Proportion of Variance	21.52 %	20.79 %	11.66 %	20.91 %	19.4 %	13.04 %	22.35 %	15.52 %	12.16 %	21.66 %	18.85 %	12.61 %
	Cum.	21.52 %	42.31 %	53.97 %	20.91 %	40.31 %	53.36 %	22.35 %	37.87 %	50.03 %	21.66 %	40.52 %	53.12 %
AC 1000T	Proportion of Variance	20.99 %	16.47 %	10.61 %	20.89 %	17.89 %	10.27 %	19.72 %	16.15 %	11.51 %	21.27 %	14.16 %	13.36 %
	Cum.	20.99 %	37.46 %	48.08 %	20.89 %	38.79 %	49.06 %	19.72 %	35.88 %	47.38 %	21.27 %	35.43 %	48.79 %

The summary table shows the proportions of explained and cumulative explained variance for the first three systematic factors of each particular liquidity measures during the period 6-13 Dec 2006. The calculation is based on projection-pursuit principal components analysis (PP-PCA) using the Croux/Ruiz (2005) algorithm. Morning period is during 10:00-11:59, the noon period is 12:00-12:50 and 13:10-14:59, the afternoon period is 15:00-16:59, and the total period is during 10:00-12:50 and 13:30-16:59

Table 9: Factor Loadings of Individual Stocks' Liquidity Measures towards their Systematic Liquidity Factors

	S [1]	S [2]	S [3]	25T [1]	25T [2]	25T [3]	100T [1]	100T [2]	100T [3]	500T [1]	500T [2]	500T [3]	1000T [1]	1000T [2]	1000T [3]
ADS	0.007	0.009	0.016	-0.002	-0.037	-0.033	0.012	0.017	-0.030	0.012	0.004	-0.032	0.004	-0.010	0.054
ALV	0.037	-0.143	0.091	0.001	-0.022	-0.057	0.011	-0.002	0.014	-0.009	-0.011	-0.050	0.036	-0.056	0.141
ALT	-0.001	-0.005	0.016	0.005	-0.019	-0.082	0.018	0.036	-0.025	0.004	-0.001	-0.007	0.005	-0.002	0.010
BAS	0.004	-0.008	0.019	0.031	-0.036	-0.012	0.011	0.010	-0.01	-0.030	-0.029	-0.017	0.010	0.001	0.012
HRX	0.018	0.012	0.001	0.029	0.035	0.042	0.002	0.016	-0.019	0.023	0.013	-0.008	0.010	-0.016	0.026
BMW	0.012	0.003	0.005	0.008	0.014	0.015	0.007	0.003	-0.015	-0.007	0.001	-0.001	-0.003	0.003	0.020
BAY	0.027	0.011	-0.008	-0.001	-0.004	-0.006	-0.007	-0.008	0.001	-0.030	-0.012	-0.028	-0.009	-0.012	0.008
CBK	0.014	-0.007	0.001	-0.012	-0.011	-0.003	-0.001	0.005	-0.008	0.001	0.000	0.002	-0.001	-0.001	0.003
CON	0.163	-0.038	0.145	0.007	-0.015	0.249	-0.020	-0.010	-0.102	-0.040	0.011	0.004	0.001	-0.046	-0.132
DCX	0.016	-0.024	0.005	0.008	0.012	0.012	-0.005	-0.006	0.001	0.019	0.015	0.025	0.001	-0.002	0.059
DBK	0.028	-0.036	0.026	-0.018	-0.029	-0.032	-0.009	-0.008	-0.001	0.059	0.041	0.078	0.045	0.031	0.110
DB1	0.550	-0.421	-0.461	-0.308	-0.308	-0.346	0.733	-0.671	-0.056	0.827	-0.552	-0.010	0.919	-0.386	0.000
DPW	0.011	-0.001	0.013	0.006	0.001	0.003	-0.005	-0.004	0.007	0.026	0.006	0.004	0.001	0.002	-0.019
DTE	0.000	0.001	0.001	-0.001	0.006	0.004	-0.002	-0.002	0.002	-0.002	0.000	-0.001	0.000	0.001	0.036
EOA	0.005	0.000	-0.013	0.173	-0.088	0.331	-0.034	-0.177	0.163	0.041	0.020	-0.024	-0.028	-0.009	-0.105
FME	0.357	0.371	-0.603	0.516	0.500	-0.625	-0.203	-0.263	0.802	0.196	0.227	0.531	-0.188	-0.502	-0.589
HEN	0.583	0.585	0.471	0.680	-0.697	-0.053	-0.640	-0.659	-0.370	-0.478	-0.750	0.412	0.196	0.438	0.222
IFX	0.000	0.002	-0.001	0.000	-0.010	-0.009	0.004	0.004	-0.001	0.006	0.000	0.001	0.001	0.000	-0.005
LIN	0.048	0.076	0.018	0.079	0.018	0.045	0.030	0.009	-0.047	0.010	0.016	-0.011	0.004	-0.033	0.070
LHA	0.007	0.005	-0.007	0.014	0.007	0.011	0.003	0.004	-0.008	0.022	0.014	0.013	0.002	0.005	-0.002
MAN	0.154	-0.028	-0.036	-0.008	-0.034	-0.111	0.043	0.032	0.001	0.007	-0.003	-0.328	0.053	0.104	0.119
MEO	-0.004	0.007	-0.012	-0.342	-0.347	-0.451	-0.016	0.013	0.000	-0.047	-0.023	-0.069	0.001	0.004	-0.001
MUV	0.009	0.026	-0.017	-0.092	-0.147	-0.170	-0.006	-0.018	0.010	0.150	0.142	0.207	0.019	-0.018	0.129
RWE	0.040	-0.034	0.020	0.013	0.010	0.012	-0.027	-0.025	0.008	-0.035	-0.024	-0.020	-0.018	-0.026	0.091
SAP	0.405	-0.555	0.405	-0.012	0.015	-0.171	-0.001	-0.003	0.113	0.099	0.237	0.611	0.263	0.602	-0.691
DPB	0.080	-0.044	0.050	0.021	0.111	0.101	-0.066	-0.097	0.402	0.036	0.011	0.075	-0.043	-0.069	-0.007
SIE	-0.012	-0.002	-0.014	-0.002	0.011	0.007	-0.009	-0.012	0.012	-0.024	-0.009	-0.020	0.002	0.032	0.039
TKA	0.010	-0.005	-0.003	0.106	-0.006	0.119	-0.013	0.011	-0.015	-0.008	-0.009	-0.014	-0.001	0.012	-0.026
TUI1	-0.002	-0.006	0.001	0.005	0.002	0.001	-0.002	-0.003	-0.003	-0.003	-0.004	0.003	-0.001	0.000	-0.001
VOW	0.029	-0.007	0.045	0.009	-0.012	-0.007	0.024	0.025	-0.020	-0.018	-0.015	0.010	-0.058	-0.137	-0.105

Factor loadings of all DAX-30 stocks towards the first three market-wide systematic liquidity factors of the five non-overlapping liquidity measures applied. The loadings are obtained by the use of PP-PCA for the time period 10:00-12:50 and 13:10-16:59 during 6-13 Dec 2006.

Table 10: Regression Statistics of Market-Wide Liquidity Factors on the Second Global Liquidity Factor

		Coef.m	Sig.m	Coef.u	Sig.u	R2.u
S	M [1]	-1.56	***	-0.44	***	0.09
	M [2]	-0.41	***	-0.41	***	0.07
	M [3]	-0.96	***	-0.60	***	0.00
AC 25T	M [1]	0.15	***	-0.06	-	0.00
	M [2]	-0.30	***	-0.05	-	0.00
	M [3]	-0.40	***	-0.16	**	0.00
AC 100T	M [1]	-0.53	***	-0.91	***	0.19
	M [2]	-0.49	***	0.36	***	0.03
	M [3]	-0.43	***	-0.12	-	0.00
AC 500T	M [1]	-2.41	***	-0.69	***	0.65
	M [2]	-1.20	***	0.09	***	0.01
	M [3]	1.90	***	0.80	***	0.30
AC 1000T	M [1]	1.19	***	0.00	-	0.00
	M [2]	-0.30	***	0.07	*	0.00
	M [3]	-3.22	***	-0.19	***	0.00
Intercept		-083	***			

Regression on the *second global liquidity factor* during the period 6-13 Dec 2006. Model statistics: Residual standard error: 0.0000; Adjusted R-squared 1.00; F-statistic: 2.71 on 15 and 2378 degrees of freedom. P-value: 0.0000.

Table 11: Regression Statistics of Market-Wide Liquidity Factors on the Third Global Liquidity Factor

		Coef.m	Sig.m	Coef.u	Sig.u	R2.u
S	M [1]	0.90	***	-0.24	***	0.04
	M [2]	1.55	***	0.62	***	0.23
	M [3]	1.27	***	0.25	***	0.03
AC 25K	M [1]	-036	***	0.11	**	0.00
	M [2]	-0.50	***	0.13	***	0.01
	M [3]	0.16	***	0.15	**	0.00
AC 100K	M [1]	0.14	***	0.04	-	0.00
	M [2]	0.56	***	0.19	***	0.01
	M [3]	0.57	***	0.36	***	0.01
AC 500K	M [1]	0.82	***	0.16	***	0.06
	M [2]	2.04	***	0.32	***	0.16
	M [3]	0.73	***	0.05	.	0.00
AC 1000K	M [1]	2.01	***	0.58	***	0.39
	M [2]	-074	***	-0.53	***	0.14
	M [3]	-0.43	***	-0.26	***	0.02
Intercept		0.63	***			

Regression on the *third global liquidity factor* during the period 6-13 Dec 2006. Model statistics: Residual standard error: 0.0000; Adjusted R-squared 1.00; F-statistic: 2.71 on 15 and 2378 degrees of freedom. P-value: 0.0000.