CFR-Working Paper NO. 09-05

Commonalities in the Order Book

H. Beltran-Lopez • P. Giot • J. Grammig

centre for financial research
Look deeper

Commonalities in the order book

Héléna Beltran-Lopez

Pierre Giot

Joachim Grammig *

January 30, 2009

Abstract

This paper uses data from one of the most important European stock markets and shows that, in line with predictions from theoretical market microstructure, a small number of latent factors captures most of the variation in stock specific order books. We show that these order book commonalities are much stronger than liquidity commonality across stocks. The result that bid and ask side as well as the visible and hidden parts of the order book exhibit quite specific dynamics is interpreted as evidence that open order book markets attract a heterogeneous trader population in terms of asset valuations and impatience. Quantifying the informational content of the extracted factors with respect to the evolution of the asset price we find that the factor information shares are highest (about ten percent) for less frequently traded stocks. We also show that the informational content of hidden orders is limited.

Keywords: limit order book, commonalities, liquidity, market microstructure

JEL classification: G10, C32

*Héléna Beltran-Lopez is a FNRS research fellow at CORE, Université catholique de Louvain, Voie du Roman Pays 34, B-1348 Louvain-la-Neuve, Belgium. Phone: +32 (0) 10 478160. Email: beltran@core.ucl.ac.be. Pierre Giot is Professor of Finance at the Department of Business Administration & CEREFIM, University of Namur, Rempart de la Vierge, 8, 5000 Namur, Belgium and CORE at Université catholique de Louvain. Phone: +32 (0) 81 724887. Email: pierre.giot@fundp.ac.be. Joachim Grammig is research fellow at the Centre for Financial Research, Cologne and Professor of Econometrics at the Faculty of Economics, University of Tübingen, Mohlstrasse 36, D-72074 Tübingen, Germany. Phone: +49 7071 29 76009. Fax: +49 29 5546. Email: joachim.grammig@uni-tuebingen.de. Send all correspondence to pierre.giot@fundp.ac.be. Work supported in part by the European Community's Human Potential Programme under contract HPRN-CT-2002-00232 Microstructure of Financial Markets in Europe and the German Research Foundation (DFG). The first author gratefully acknowledges financial support from the Belgian 'Fonds National de la Recherche Scientifique' (FNRS). We are grateful to Deutsche Böerse AG for providing access to the limit order data. Earlier drafts of the paper were presented at seminars at CREST and Frankfurt University. We thank the seminar participants, Michel Baes, Frank de Jong and Thierry Foucault for helpful comments. We owe special thanks to Stefan Frey for his collaboration in the reconstruction of the limit order books, and to Kerstin Kehrle for her excellent research assistance. Uwe Schweickert provided his invaluable expertise about the details of the Xetra trading system. The usual disclaimer applies.

1. Introduction

More and more trading venues throughout the world operate as open order book markets. In those exchanges, liquidity is supplied voluntarily by market participants who provide an inflow of limit buy and sell orders. Non-executed orders constitute the limit order book which consists of distinct, sorted limit price-depth pairs. Since in a pure open order book market the book is the single source of liquidity, identifying the factors that account for its variation is of paramount interest for regulators, exchange operators and investors.

This paper studies commonalities of liquidity supply in a pure limit order book market. We draw on results from microstructure theory that hint at the existence of common factors which explain the inter-temporal variation of an order book. According to Glosten's (1994) seminal paper, limit order traders post orders with characteristics (price and volume) that depend on the underlying asset value. In Glosten's framework, limit order traders protect themselves against the picking-off risk, i.e. adverse price changes in the value of the security. Accordingly, a single latent factor - the underlying asset value - affects all price-depth pairs displayed in the book. The papers by Parlour (1998) and Foucault, Kadan, and Kandel (2003) suggest that another factor accounts for the variation of the order book: the mix of patient and impatient traders. In their theoretical frameworks, just like in real world trading systems, orders are executed according to time and price priorities. Traders thus face a trade-off between the price improvement made possible by a limit order and faster execution using a market order or an aggressive limit order which jumps ahead of the order queue. Impatient traders incur higher waiting costs, hence they are more likely to post market orders, or aggressive limit orders that improve the best quotes. Patient traders are ready to provide liquidity using less aggressive limit orders away from the best quotes. Accordingly, the shape of the order book depends on the mix of patient and impatient traders. This paper aims to provide empirical evidence on the existence of commonalities in the order book. For that purpose we address the following research agenda:

- Is there evidence that the price-depth pairs in a limit order book exhibit those commonalities suggested by microstructure theory? In other words, can liquidity supplied in a limit order book be 'summarized' by a small number of latent factors, and if yes, what is their specific role in explaining the variations of the book?

- The theoretical papers cited above implicitly assume some degree of symmetry between the buy and sell sides of the order book. However, if agents with different liquidity needs, asset valuations, and risk aversion place orders on both sides of the book, different factors may explain the variation of the buy and sell sides of the order book. Accordingly, this paper investigates whether ask and bid sides of the book feature the same commonalities. In the same vein, since trader heterogeneity may be reflected in the use of visible or hidden limit orders, we also analyze whether the visible and the hidden part of the order book exhibit the same commonalities.

- Recent empirical papers (e.g. Cao, Hansch, and Wang (2004)) show that the limit order book contains informational content with respect to the evolution of the fundamental asset price. Relating our paper to this literature, we assess the informational content of the extracted factors, especially the differences between small, less frequently traded and large, actively traded stocks. In the same vein, we also quantify the informational content of the hidden part of the book.

To a reader familiar with the recent literature on commonalities in market microstructure, this research agenda may appear surprising. As a matter of fact, most of the previous work has focussed on liquidity commonalities across stocks. Recent papers along this line of research include Chordia, Roll, and Subrahmanyam (2001), Hasbrouck and Seppi (2001), Brockman and Chung (2002), Hansch (2003), Bauer (2004), Coughenour and Saad (2004), Domowitz, Hansch, and Wang (2005) and Kempf and Mayston (2006). In contrast, we focus on the commonalities exhibited by price-depth pairs in stock specific order books. However, in order to link this paper to the previous literature, we use the results of the stock specific analysis and also quantify cross-sectional commonalities of liquidity.

Our empirical analysis is based on Frankfurt Stock Exchange (FSE) order book data from the Xetra trading system. This unique dataset is particularly well suited for the analysis of the present paper. The data contain the complete order books of the stocks constituting the DAX30 index for a three months period in 2004.[1] Similar to Euronext, the other prominent European trading system, Xetra is an automated auction system with an open limit order

[1]The DAX30 is a value weighted index of the thirty largest German stocks and one of the leading European stock indexes, along with the French CAC and the British FTSE.

book.[2] There are no designated market makers for frequently traded stocks. Xetra is a pure open limit order book market. In contrast to the NYSE or NASDAQ, the Xetra order book does not compete with significant other purveyors of liquidity. Although a floor trading system and some regional exchanges offer alternatives to the Xetra book, they only account for a small amount (2% to 5%) of the total trading activity. Some of the stocks in our sample are also cross-listed at foreign trading venues, most importantly at the NYSE. Daimler-Chrysler is the most prominent example. However, compared to FSE/Xetra, the foreign trading volume is quite small. The data used in the paper originate directly from the Xetra production system. They include all necessary information to perform a real time reconstruction of the stock specific order books over the sample period. Since the FSE itself uses these data for internal and external reporting, the data quality is excellent.

The empirical methodology employed in this paper tries to keep things simple. Using sequences of reconstructed real-time order books we compute price impacts which measure the relative price per share increase (for a buyer initiated transaction) or decrease (for a seller initiated transaction) when hypothetically trading a volume v at time t against the book. Accordingly, price impact curves monotonically increase (buy side) or decrease (sell side) with the hypothetical trading volume. They represent a natural way to measure committed liquidity supply in an open order book system. Performing a principal components analysis (PCA) on the price impact data we extract orthogonal factors that account for the variation of liquidity supply in an open order book.[3] The extracted factors and the factor weights are then further analyzed to address the research agenda outlined above.

The main results can be summarized as follows: For diurnally adjusted price impacts, two principal components already provide an explanatory power of more than 94% (in terms of cumulative R^2). This holds true for the 'one-sided' analysis, i.e. when the PCA is performed separately for the buy and the sell side of the order book. Variations in the first principal component shift the supply and demand curves while the second principal component rotates the slope of the book. We show that bid and ask side of the book are driven by distinct, albeit correlated factors. When performing PCAs on both sides of the book simultaneously, four factors are needed to achieve a similar explanatory power as in the "one-sided" analysis.

[2]The Xetra system also operates at the Dublin and Vienna stock exchanges and at the European Energy Exchange.

[3]Like volatility and inside spreads, the price impact series exhibit a intra-day seasonal pattern (diurnality), hence the price impact series are diurnally adjusted prior to performing the PCA.

Similarly, we show that the visible and hidden portions of the limit order book share some common dynamics, but also exhibit clear idiosyncracies. We interpret these results as evidence that an open order book market attracts a heterogeneous population of limit order traders with different trading strategies, trading needs, and asset valuations. Our results also indicate that while there is evidence of liquidity commonality across stocks, the total explanatory power of the principal components is much smaller than for the stock specific analysis. Our cross sectional commonality results are, however, broadly comparable to those reported in Hasbrouck and Seppi (2001) and Bauer (2004).

The paper also shows that the information share attributable to the extracted factors with respect to the long run evolution of the asset price is non-negligible. In other words, shifts and rotations of the order book carry informational content. The information shares are considerably different across stocks. While for the group of most actively traded stocks (which are also the biggest in terms of market capitalization) we estimate an average information share attributable to the extracted factors of about 5 percent, the number doubles for the group of least frequently traded stocks. On the other hand, the hidden part of the book does not carry economically significant informational content.

The remainder of the paper is organized as follows. Section 2 outlines the market structure and conducts a descriptive analysis of the data. In Section 3 we review the empirical methodology. Sections 4.1 (one sided analysis), 4.2 (joint analysis of ask and bid side), 4.3 (joint analysis of visible and hidden book), and 4.4 (analysis of incremental price impacts) discuss the results of the principal component analysis performed on stock specific reconstructed order books. Section 4.5 focusses on cross sectional commonalities. Section 4.6 sheds light on the informational content of the extracted principal components. We conclude in Section 5 with a summary and an outlook for further research.

2. Market structure and order book liquidity measures

2.1. The Xetra order book

Xetra is an open order book system developed and maintained by the German Stock Exchange.[4] It operates since 1997 as the main trading platform for German blue chip stocks

[4]See Deutsche Börse AG (1999) for a detailed description of the Xetra system.

at the Frankfurt Stock Exchange. Between an opening and a closing call auction - and interrupted by another mid-day call auction - trading is based on a continuous double auction mechanism with automatic matching of orders based on rules of price and time priority. During pre- and post-trading hours it is possible to enter, revise and cancel orders, but order executions are not conducted, even if possible. Trading hours extend from 9 a.m. CET to 5.30 p.m. CET. For the DAX30 stocks there are no designated market makers. Traders can view the full order book, except hidden shares coming from so-called iceberg orders. An iceberg order is similar to a limit order in that it has pre-specified limit price and volume. The difference is that a portion of the volume is kept hidden from the other traders. The visible portion of the iceberg order, called the 'peak', enjoys full price and time priorities as any visible limit order. The hidden portion receives only price priority. All disclosed volumes are executed first, even if those volumes entered the book after the iceberg order submission. When a market order hits the hidden portion, a new portion of the iceberg order (equal to the peak size) is revealed to the market participants and is granted time priority over subsequent order submissions. Consequently, a trader submitting a market order may receive an unexpected price improvement if her market order is executed against a hidden order.

The Xetra trading protocol is comparable to the Euronext trading system (which, at the time of this writing, encompasses the Amsterdam, Brussels, Lisbon and Paris stock exchanges). Xetra also shares many features with the trading system of the Hong Kong stock exchange, which has been recently studied by Ahn and Cheung (1999), Brockman and Chung (1999) and Ahn, Bae, and Chan (2001). A feature which makes the Xetra data especially useful for the purpose of our study is that market orders exceeding the volume at the best quote 'walk up the book'. In other words, the system guarantees market orders immediate full execution, at the cost of a worse price per share than the best quote. The liquidity measures used in the present paper implicitly assume that such a 'walk up the book' is possible.

The German Stock Exchange granted us access to a database containing complete information about Xetra order book events (entries, cancelations, revisions, expirations, partial-fills and full-fills of market, limit, and iceberg orders) that occurred at the FSE during the first three months of 2004 (January, 2nd - March, 31st). The data encompasses the 30 stocks belonging to the DAX30 index. Based on the time stamped event based data we perform a real time reconstruction of the order book sequences. Specifically, starting from an initial set of non executed limit orders, we track each change in the order book implied by entry, partial or full-fill, cancelation and expiration of market, limit, and iceberg orders.[5] The result is a sequence of order books, i.e. price-depth pairs which are sorted in price descending (price ascending) order for the bid (ask) side. From these order book sequences, snapshots at 5-minute intervals during the continuous trading hours are taken. This results in 102 order book snapshots per day. To account for the presence of iceberg orders, which are not fully disclosed, we actually track two limit order book sequences. First, we have the visible order book, which contains all limit orders as well as the visible portion of the iceberg orders. Second, we reconstruct the complete order book that features all orders, including the hidden portions.

Table 1 presents some descriptive characteristics of the data. One can see that liquidity supply and demand is quite active. On average, 13,000 (11,000) non-marketable limit orders per stock are submitted (canceled) per day. Implicit transaction costs are relatively small with relative spreads ranging from 0.04% to 0.14%. On average, 2,100 trades per stock are executed each day and 15.2% of these walk up the book (i.e. they are matched by standing limit orders beyond the best bid and ask prices). We refer to those events as 'aggressive trades'.

2.2. Using price impacts to measure liquidity beyond the best quotes

Automated auction markets enforce price and time priority rules which govern the trading process. This implies that the price impact of a buy (sell) market order is an increasing (decreasing) function of the trade size. The available pre-trade liquidity of the book can be characterized by the unit price for selling v shares at time t,

$$b_t(v) = \frac{\sum_k b_{k,t} v_{k,t}}{v} \tag{1}$$

where v is the volume executed at k different unique bid prices $b_{k,t}$ with corresponding volumes $v_{k,t}$ standing in the limit order book at time t. The unit price $a_t(v)$ of a buy of size v at time t can be computed analogously. The unit prices $b_t(v)$ and $a_t(v)$ can be computed for an arbitrary range of volumes v. Liquidity supply in an order book can thus be characterized

[5] This is done by implementing the rules of the Xetra trading protocol in a GAUSS program. An exhaustive battery of consistency checks showed that there were no errors during the reconstruction process.

at any given point at time t by a grid of unit ask and bid prices conditional on the traded volume v. Three transformations to the raw unit price series $a_t(v)$ and $b_t(v)$ are needed to make the data amenable for our empirical analysis.

First, the unit price series will be non-stationary, but our statistical methodology requires stationary data input. We therefore consider a stationary transformation into "price impacts". Price impacts capture the premium (discount) incurred by a buy (sell) market order trader when the initiated transaction is executed against standing limit orders beyond the best quotes. Specifically, bid price impacts are computed as: [6]

$$bp_t(v) = \frac{b_t(1) - b_t(v)}{b_t(1)} \cdot 100. \tag{2}$$

Ask price impacts, $ap_t(v)$, are computed accordingly.

Second, to ensure comparability across stocks, we do not express the trade size v in number of shares, but in euros. By construction, the larger a seller (buyer) initiated trade against the book, the lower (higher) the per share price paid as the market order hits more and more limit orders and is likely to walk up further in the book. For example, a $bp_t(100,000) = 1\%$ bid price impact means that the per share price the initiator of a 100,000 euros sell order at time t receives for each share sold is 1% lower than the price if the seller's volume would be smaller or equal to the depth at the best bid. When the trading volume v does not consume all the shares available at the best price, the price impact is zero.

Third, we account for diurnality (intra-day seasonality) of the price impact series. Figure 4 shows that the intra-day pattern is characterized by high price impacts in the morning, a stablization during the day, and a slight increase at the close. Adopting the standard procedure in the recent empirical literature, we remove intra-day seasonality by subtracting time-of-day (tod) means and divide by the time-of-day standard deviations. To compute tod means and standard deviations the trading day is divided into thirty-minute bins. Specifically, diurnally adjusted bid price impacts $\widetilde{bp}_t(v)$ are computed as

$$\widetilde{bp}_t(v) = (bp_t(v) - bp_{tod}(v))/std(bp_{tod}(v)), \tag{3}$$

[6] See Irvine, Benston, and Kandel (2000) and Gomber, Schweickert, and Theissen (2004) for a similar measure of book liquidity. Note that in microstructure the term 'price impact' is also used in a different context. The difference of effective and realized spread is often also referred to as price impact.

where $bp_{tod}(v)$ is the average and $std(bp_{tod}(v))$ the standard deviation of the price impact in the corresponding thirty minute time-of-day bin. An analogous formula is used to compute diurnally adjusted ask price impacts.

We compute ask and bid price impacts for a grid of trade sizes v ranging from 10,000 euros to 500,000 euros, with an increment of 10,000 euros. This yields a total of 50 price impacts for both the ask and bid sides. The choice of the minimum and maximum volumes is based on the following rationale. We have computed (for the 30 stocks) the quantiles of the market orders and marketable limit orders submitted during the time period under analysis. These empirical results are summarized in Table 2 (we report results averaged across the 30 stocks). The median trade size (on average) is approximately equal to 23,000 euros and the 90% percentile is at 100,000 euros. While the 99% percentile is a bit larger than 300,000 euros, the maximum traded volume is larger than 3 millions euros. The information conveyed by the 99% percentile indicates that extremely large volumes are almost never observed. Besides, it is well known that traders strategically time their large orders, i.e. they enter large order when liquidity is plentiful in the order book. Therefore, it seems not sensible consider extreme trade sizes. A stock specific analysis reveals that this grid is adequate for all stocks in the sample. In a previous version of this paper, we have considered a grid size ranging from 50,000 euros to 2 million euros, with an increment of 50,000 euros. The results are qualitatively similar, hence the analysis is robust to the choice of the grid size.

To further illustrate, Table 3 reports the average price impacts and standard deviations for hypothetical trade sizes of 20,000 (small trade), 50,000 (medium trade), 100,000 (large trade) and 500,000 euros (extremely large trade). For small and medium trades, the price impacts are small (around 0.01% and 0.02% respectively). Even for relatively large trades, the provision of liquidity is quite adequate: A 100,000-euro market order only incurred a 0.03% price impact. Extremely large orders, however, had to walk up the book extensively, resulting in larger price impacts.

A key feature of the data is that we have information about both the visible and hidden orders present in the book. This implies that, during the reconstruction process, we can keep track of two order book sequences. The first only contains the shares visible to traders (the 'visible' order book), whereas the second also includes the hidden orders (the 'complete order book'). Consequently, three types of price impacts can be computed. We refer to *visible* price impacts as those that results from an application of Equation (1) using only the visible book depth, while *complete* price impacts are computed using the complete book, i.e. including hidden volumes. The latter are, by definition, smaller than or equal to the former. Third, we compute "hidden price impacts". The hidden price impact measures the price improvement brought about by the hidden part of the iceberg orders. Hidden price impacts are computed as the difference between the price impact implied by the complete order book and the price impact implied by the visible order book.

2.3. Time series properties of price impacts

The time series properties of the price impacts are remarkably similar for the 30 stocks in our sample. Hence we present our results as holding 'on average for all stocks' and present averages across stocks along with confidence bounds indicating the distribution across stocks. For illustration purposes we also present the results for one single stock. For no particular reason, except that it can be considered a "typical" stock, we chose Bayer AG (ticker symbol BAY), a company specialized in life sciences. Figure 1 presents the time series of 20,000-, 50,000-, 100,000- and 500,000-euro ask price impacts for the Bayer stock. The price impacts were computed using the complete order book. The time series look similar for both the bid side price impacts and the price impacts computed from the visible book. Figure 1 and Table 3 indicate that small volume price impacts exhibit little persistence and a high dispersion index. Figures 1 and 2 show that for relatively small volumes the price impacts exhibit only small serial correlation, while the 500,000-euro price impacts are strongly autocorrelated. All series show intra-day seasonality, although the diurnal pattern is less pronounced for the smallest price impacts. Once diurnally adjusted, the series exhibit very similar dynamics, except that there are no seasonal peaks left in the autocorrelograms (see Figure 3). Overall, these results indicate that liquidity supply is far from being 'static' which corroborates previous evidence reported by Ahn, Bae, and Chan (2001) and Biais, Hillion, and Spatt (1995) for the Hong-Kong and Paris Bourse stock exchanges, respectively.

Table 5 reports the cross-correlations between the inside spread and the price impacts. The first six rows of this table report the spread correlations with ask and bid price impacts for volumes ranging from 20,000 to 500,000 euros. The reported correlations are very small from an economic perspective. Indeed, they are often almost equal to zero, which suggests that the inside spread exhibits other dynamics than the price impacts. This shows that the dynamics of the full book liquidity (the focus of our work) is distinct from the near-the-best-quotes liquidity. Table 5 reports correlations of the inside spread with the Xetra Liquidity Measure (XLM), the official liquidity measure computed by the FSE (see Gomber, Schweickert, and Theissen (2004)). The XLM is constructed in a similar way as our price impacts. However, while we compute price impacts relative to the best bid or ask, the XLM standardizes volume-dependent per share prices with respect to the the mid-quote (the average of best ask and bid). The XLM thus captures, by construction, not only the variation of book liquidity supply beyond the best quotes, but also that of the spread. Not surprisingly, the XLM is much more correlated with the spread than the price impacts. The XLM is thus not well-suited for the analysis of the dynamics of the book liquidity beyond the best quotes. Finally, the high cross-correlations reported in Table 4 indicate that price impacts for different trade sizes indeed display common dynamics. This result provides the springboard for the analysis in the next section.

3. Modeling the order book using principal component analysis

Several statistical methodologies are at hand to study order book commonalities. In this paper we employ a standard technique, principal component analysis (PCA).[7] This section describes the PCA methodology in the context of order book modeling. A reader familiar with the methodology can jump ahead to Section 4 and use this section as a reference for notational details.

The data input for the PCA are the diurnally adjusted time series of price impacts $\widetilde{ap}_t(v)$ and $\widetilde{bp}_t(v)$. We compute these price impacts for a grid of N trading volumes and for a sequence of T equidistant order book snapshots during the three-month sample period. The price impacts are collected in a $T \times N$ matrix X such that each column of X contains a time series sequence of price impacts computed for a given volume v. For a one-sided analysis the columns are sorted by increasing v.

We also conduct a joint analysis on horizontally concatenated price impacts from both sides of the book, and joined hidden and visible price impacts. Hence, we need a general

[7]In an earlier draft we have also conducted a dynamic factor analysis as advocated by Forni, Hallin, Lippi, and Reichlin (2003). However, we found that for the context of the present paper the dynamic factor analysis offers no additional benefit, but comes at the cost of a higher degree of complexity.

notation to indicate the column of the input data matrix. For that purpose, let $x_{t,i}$ denote the i-th column price impact at snapshot time t. The row index $t = 1, \ldots, T$ identifies the time of the order book snapshots. Since PCA requires mean zero, one standard deviation data input, the price impact data are standardized as follows:

$$z_{t,i} = (x_{t,i} - x_i)/sdx_i. \tag{4}$$

x_i and sdx_i denote sample mean and standard deviation of column i, respectively. The standardized price impacts $z_{t,i}$ are collected in a $T \times N$ matrix Z.

The standardized price impacts are expressed as a linear combination of N orthogonal vectors, referred to as principal components,

$$Z = P\,W', \tag{5}$$

where $P = (F1, F2, \ldots, FN)$ denotes a $T \times N$ matrix containing the orthogonal principal components ('factors'), and W is a $N \times N$ matrix of weights. The i-th $(i = 1, \ldots, N)$ column of P is referred to as the i-th principal component, and denoted $Fi = (Fi_1, Fi_2, \ldots, Fi_T)'$.[8] The extraction of orthogonal principal components is achieved by computing the eigenvalues and associated eigenvectors of the sample correlation matrix of standardized price impacts, $\frac{1}{T}Z'Z$. The columns of the weight matrix W containing these eigenvectors are arranged in descending order of the associated eigenvalues. Finally, the principal components are computed as $P = Z\,W$. The weights/eigenvectors in W provide the mapping between the principal components and the price impact series. We will show below how an analysis of the weight matrix W helps understand how the principal components 'explain' the variation of the shape of the order book.

Ordering by decreasing eigenvalues implies that the first principal component, $F1$, is the one with the largest variance, i.e. it explains most of the variation in the price impacts. The second principal component, $F2$, has the second largest variance and explanatory power and so forth. The explanatory power (referred to as cumulative R^2) of the first n principal components $F1, F2, \ldots, Fn$ can be computed by dividing the sum of the n first eigenvalues

[8] When the PCA is performed on ask and bid side price impacts, we denote the principal components FAi and FBi, respectively.

by the total sum of all N eigenvalues.

If the price impact series share common components (exhibit commonalities), then the first $n << N$ principal components will have a considerable explanatory power. The first n principal components, collected in the $T \times n$ matrix $P^{(n)} = (F1, F2, \ldots, Fn)$, combined with the weight matrix $W^{(n)}$ (which contains the first n columns of W) then feature enough information to 'rebuild' the whole state of the order book at time t. For that purpose we first compute $\widehat{Z} = P^{(n)}W^{(n)}$ and then re-introduce the standard deviation and mean of each column using the inverse transformation of Equation (4). This provides $\widehat{X}$, the reconstructed price impacts matrix derived from the information provided by the first n principal components. If the price impacts feature commonalities then the reconstructed price impact series will be highly correlated with the observed price impacts. Provided that n is small, the PCA can deliver a considerable reduction of complexity. In our application this would imply that the variation of the order book can be explained by a small number of orthogonal factors.

4. Empirical evidence on order book commonalities: Results and discussion

In this section we perform PCAs on stock specific price impact data, first on each side of the book separately (Section 4.1), before performing joint PCAs for both the sides of the order book (Section 4.2) and for visible and hidden impacts (Section 4.3). Incremental price impacts are analyzed in Section 4.4. Cross sectional commonalities are investigated in Section 4.5. Since the weight structure and the explanatory power of the significant principal components are remarkably similar for all stocks, the tables and figures which document our results report averages across stocks.[9] Whenever useful, we also report cross sectional standard deviations, minima and maxima as well as confidence bounds.

4.1. One-sided analysis of commonalities in the order book

We start with the visible order book and perform a principal components analysis on the diurnally adjusted price impact series using the grid of 50 volumes from $v = 10,000$ through

[9]Note that this does not imply that the order book of the 30 stocks are generated by the same factors. Indeed, the factors for the 30 stocks are different, although they may share some common dynamics. We deal with this issue in Section 4.5

$500,000$ euros. The PCA is conducted for ask and bid side price impacts separately. The results reported in Table 6 show that the first two principal components explain (in terms of cumulative R^2) more than 94% of the variation in visible price impacts. This suggests a choice of $n = 2$ principal components, a decision which is also supported by the low explanatory power of the third principal component (around 3%). There are only minor differences between bid and ask side. Figure 5 depicts the weights of the first two principal components for the bid and the ask side. The weights of the first principal component are positive and approximately the same for volumes ranging from 100,000 to 500,000 euros. Weights decrease by about 50% when v decreases from 100,000 euros to 10,000 euros. An increase in the first principal component therefore produces an upward shift in the whole price impact curve, which, however, affects the price impacts for volumes smaller than 100,000 euros to a lesser extend. This is illustrated in the top panel of Figure 7 where we perform simulations using the Bayer (BAY) PCA results as an example. Starting from a base scenario, we successively increase and decrease the value of the first principal component and plot the change in the whole range of price impacts. The top panel of Figure 7 shows how the ask side of the order book shifts upwards (downwards) when the first principal component is increased (decreased). It is clearly visible that the price impacts of large volumes v are more affected than the price impacts of small volumes. The top panel of Figure 7 suggests referring to the first principal component as 'shift factor' since it moves all price impacts in the same direction, albeit in a non-linear fashion: An increase of the shift factor indicates that liquidity quality is mildly reduced near the inside quotes, while beyond medium size trading volumes liquidity quality deteriorates more sharply.

Figure 5 shows that the weights of the second principal component are negative for volumes up to about 230,000 euros and positive for larger volumes. An increase in the second principal component thus reduces the price impacts for smaller volumes, and increases those for large volumes. This effect is illustrated in the bottom panel of Figure 7 where the ask side of the Bayer order book is again used as an illustrative example. As before, we start from a base scenario, increase and decrease the value of the second factor and plot the resulting price impacts. Figure 7 shows that, as the second principal component is increased, the whole ask side of the book rotates counterclockwise around a volume of about 230,000 euros. Price impacts for low through medium volumes decrease, while price impacts increase for medium to very large volumes. Put differently, an increase (decrease) in the second principal component indicates an improvement (deterioration) of liquidity quality for small to medium-volume trades, while liquidity beyond medium-volume trades is negatively (positively) affected. It seems natural to refer to the second principal component as 'rotation factor'.

Figure 5 shows that the weight patterns are remarkably similar for both sides of the book. We come back to this issue below when we investigate whether the ask and bid sides are subject to the same dynamics. Additional results from PCAs performed on the complete order book (including hidden price impacts) are presented in Table 6. The reported figures for the explanatory power and the pattern of the weight structures (not reported) are very close to what we find for the visible order book.

4.2. Joint analysis of ask and bid sides

The 'one-sided' analysis in the previous subsection revealed that the first two principal components (shift and rotation factor) have a remarkable explanatory power for both sides of the order book, and that the ask and bid side weight structures are quite similar. It could therefore be conjectured that one would also need only two factors (one to account for the shift and one for the rotation of the price impact curves) when performing a joint PCA using the price impacts from sides of the book. However, the empirical results reported in Table 6 contradict this hypothesis. These results rather point to asymmetries in the order book.

The bottom panel of Table 6 reports the explanatory power of the first four principal components extracted by the joint PCA. Averaged across stocks, the cumulative R^2 of the first two principal components amounts to 82%, a considerable reduction compared to the cumulative R^2 of 94% from the one-sided analysis. The explanatory power of the three first principal components amounts to 88%, and it takes four principal components to reach a cumulative R^2 of 94%. This suggests that different factors account for the variation of the bid and ask side of the order book.

The interpretation of the four 'joint' principal components is not as clear-cut as in the single sided PCA and warrants some further investigation. For that purpose we show in Table 6 that the explanatory power of the first two joint factors is quite similar to the explanatory power of the first principal component (shift factor) from the one-sided analysis (about 82%). This suggests that the one-sided shift factor captures the same dynamics as the first two principal components from the joint PCA. If this bears out, then price impacts reconstructed from the shift factor should be closely correlated with price impacts reconstructed from the first two joint principal components. An analogous reasoning can be applied with respect to the one-sided rotation factor and the third and fourth principal components from the joint PCA. Indeed, the correlations between the price impacts reconstructed from the shift factor and the price impacts reconstructed from the first two joint factors are greater than 0.98 (for all price impacts and all stocks, both for the bid or the ask side). This shows that the first two factors from the joint PCA indeed capture the same order book variation as the shift factor from the one-sided analysis. Similar results are obtained regarding the relation of the one-sided rotation factor to the third and fourth principal components.

Further evidence on the asymmetries between the bid and ask sides of the book is provided in Table 4. Here we report the cross-market side correlations of the raw price impacts for different trade sizes. All correlations between the bid and ask side price impacts of the visible book are smaller than 0.25. Furthermore, the cross-market side correlations between shift and and rotation factor are small. Specifically, opposite market side shift factors are only weakly correlated (about 0.13 averaged across all stocks), and the correlation between the opposite side rotation factors is negligible. Ask side shift factor and the bid side rotation factor are also only weakly correlated (about 0.07 on average). The same holds true for the correlation of the bid side shift and ask side rotation factors.

As a matter of fact, the apparent heterogeneity of the buy and sell sides of the order book is not surprising. In limit order markets like Xetra or Euronext, limit order traders have no compelling reason to be simultaneously active on both sides of the market. An open limit order market is in that respect quite different from a dealership system where dedicated market makers have a binding obligation to supply liquidity for both ask and bid side in order to maintain an orderly market. The classic inventory model of Stoll (1978) assumes such a market structure. Stoll's model implies symmetrical ask and bid side variations as the dealer shifts her bid and ask prices simultaneously to manage her inventory position. Our results indicate that an anonymous open order book market attracts heterogeneous liquidity suppliers. The heterogeneity can refer to asset valuation, impatience and being informed about the true asset value. In an analysis of the Reuters D2000-2 order book trading system Danielsson and Payne (2001) reach the same conclusion.

4.3. Commonalities of visible and hidden liquidity

In this section the issue of trader heterogeneity is further investigated by a study of commonalities between the visible and hidden liquidity contained in the order book. For that purpose the PCAs are performed on horizontally concatenated data sets which contain both visible and hidden ask price impacts. As described in Section 2.2, the hidden price impacts capture the price improvements brought about by iceberg orders. The PCAs are performed one-sided, i.e. separately for the buy and the sell side. As before, all price impacts are diurnally adjusted.

Table 6 shows that four principal components are needed to achieve an explanatory power comparable to the one-sided PCA performed on visible price impacts. While in the latter application shift and rotation factor deliver a cumulative R^2 of about 94%, the explanatory power of the first two principal components in the joint visible-hidden PCA is reduced to 80 % (averaged across stocks). We have to include the third and fourth principal components to increase the cumulative R^2 to 94%.

Performing a separate PCA on the hidden price impacts only, the results are qualitatively similar to the results for the visible book.[10] Specifically, the first two principal components explain more than 90% of the variation of the hidden price impacts, and the interpretation of the first principal component as 'shift factor' and the second as 'rotation factor' also applies. Yet, the variation of the hidden part of the book is driven by other factors than the visible part. The correlation between the (same side) hidden book shift factor and the visible book shift factor is weak, about 0.21 (cross stock averages). The correlation between the visible and hidden rotation factor is about the same size. The trader heterogeneity story of the previous subsection thus extends to the provision of visible and hidden liquidity.

4.4. Commonalities in incremental price impacts

The analysis in Section 4.1 focused on the price impacts in levels. However, price impacts for two volumes are, by construction, correlated. For example, the ask price impact of a $v = 10,000$ euros trade and the ask price impact of a 11,000 euros trade share the price impact of the first 10,000 euros. To check the robustness of our conclusions we therefore check whether price impact differences exhibit the same commonalities. For that purpose we

[10] To conserve space, we do not report tables with the results of this analysis.

compute the differences of two adjacent price impacts and perform the PCA on these data. For example, the price impact difference at 20,000 euros is computed as the price impact at 20,000 euros less the price impact at 10,000 euros. These price impact differences are constructed for all volumes in the grid up to 500,000 euros. The principal component analysis of the previous sections is then performed on price impact differences for the visible, complete and hidden order books. The PCA results are presented in Table 7. They show that the conclusions of Sections 4.1, 4.2 and 4.3 extend to price impact differences, too. Although the explanatory power of the first two principal components extracted from one-sided PCAs (67% - 70%) is lower than for the level analysis, it is still remarkable and shows that price impact differences, too, exhibit strong commonalities. The weights for the first and second principal components are depicted in Figure 6. This is the counterpart of Figure 5 which shows the level price impact results. The plots are very similar, such that the interpretation of the first two principal components as shift and rotation factors also applies to price impact increments.

4.5. Cross-sectional commonalities

As mentioned in the introduction, the analysis performed in the previous subsections is quite different from the 'usual' commonality analysis studied in the market microstructure literature. Hasbrouck and Seppi (2001) and Coughenour and Saad (2004) (for the NYSE), and more recently Hansch (2003), Bauer (2004) and Kempf and Mayston (2006) (for automated auction markets) use PCAs to study liquidity commonalities across stocks. In order to link our paper to that literature, this section performs a cross sectional principal component analysis of liquidity-related variables for the 30 stocks in our sample. To perform an analysis along the lines of Hasbrouck and Seppi (2001) one has to select a single liquidity variable per stock upon which the the cross sectional PCA can be performed. In Table 8 we reports the results for a whole set of alternative liquidity indicators. It is convenient to draw on the stock specific results obtained before, and to use the previously extracted ask and bid side shift and rotation factors. These are obtained from one-sided principal components analysis of the visible price impacts (Section 4.1). As discussed above, the first two PCA extracted factors summarize stock specific variations of liquidity, and thus are obvious choices to study commonalities across stocks. We also use the ask side price impacts for 10,000 euros and

100,000 euros trades. By performing the PCAs on price impacts for two specific trade sizes, we check whether there is more commonality within similar sections of the book across stocks than for the whole book. Finally, we also use the inside spread and the best ask depth (including hidden liquidity). While the PCAs performed on shift and rotation factors and the price impacts deliver information about commonalities in book liquidity (i.e. liquidity supply for aggressive trades), the cross sectional PCAs performed on the inside spread and best ask depth help measure the amount of commonality in the provision of liquidity at the best quotes (for small trades).

Table 8 shows that there is obviously much less commonality across stocks than in the books for individual stocks. Choosing the ask side shift factor as liquidity indicator, the cross-sectional commonality is around 18% (cumulative R^2 of the first three principal components). Using the ask side rotation factor commonality measured in the same way amounts to 13%. The cross-sectional analysis of the order book variables (ask 10,000 euros and 100,000 euros price impacts as well as spread) broadly yields the same results. We find almost no commonality for the best ask depth. The results are similar for the bid side.

These results show that cross sectional liquidity commonality is much smaller than commonality in stock specific order books. Our cross sectional results are, however, broadly comparable with those in Hasbrouck and Seppi (2001) who report about 30% explanatory power for the first three principal components in their cross sectional commonality study. In a more recent study Bauer (2004) reports about 25% commonality for the Swiss stock market, which also operates as an open order book market. The result that stocks traded in open order book systems exhibit less commonality in liquidity is a result that warrants further analysis.

4.6. The informational content of the factors

In this section we address the following questions: Can the previously identified principal components help explain long-run movements of the fundamental asset price proxied by the mid-quote? In the affirmative, which of the principal components (the shift factor or the rotation factor) is the most informative? Does the hidden portion of the book provide information beyond what is already contained in the visible book?

Ours is not the first attempt to address those issues. As a matter of fact this section provides, from a different methodological angle, a contribution to a growing empirical literature focusing on the informational content of the order book (Pardo and Pascual (2004); Cao, Hansch, and Wang (2004); Naes and Skjeltorp (2005)). In this paper we address these questions using a standard time series methodology. The basic idea is to estimate a vector autoregressive (VAR) system that contains the mid-quote change and the principal components as endogenous variables. We then analyze the long-run mean square forecast error of the mid-quote change. This variance decomposition gives the proportions of the movements of the mid-quote changes attributable to the principal components. In finance the resulting variance shares are often referred to as information shares, Hasbrouck (1991) is the classic reference.[11]

In order to assign variance shares to the orthogonal innovations of each endogenous variable, the contemporaneous correlation of the innovations of the VAR must be disentangled. The standard solution is to perform a Cholesky decomposition of the covariance matrix of the estimated VAR residuals.[12] Depending on the degree of contemporaneous correlation, a change in the Cholesky ordering will result in different information shares. To assess the sensitivity of our results we thus vary the ordering of mid-quote change and principal components. One ordering places the mid-quote change first, which amounts to assuming that the shape of the book does not contemporaneously affect the mid-quote. This ordering tends to reduce the information share assigned to the principal components. Alternatively, we place the mid-quote last and put the principal components first. This tends to increase the information share of the factors. The alternative orderings provide upper and lower bounds of the informational content of the factors. We also vary the ordering of the principal components by interchanging the position of shift and the rotation factor.

We focus on two VAR specifications. In the first, the endogenous variables are the shift and rotation factors obtained from separate stock specific PCAs performed on ask and bid side visible price impacts. We refer to this specification as the 'visible factor VAR'. The second VAR includes the factors extracted from a PCA performed on both visible price impacts and incremental hidden price impacts. This specification - which we refer to as 'visible and hidden

[11] We assume that the reader is familiar with the basics of structural VAR analysis and innovation accounting via long run MSE decompositions. Chapter 11 in Hamilton (1994) provides a lucid review of the methodology.

[12] This amounts to assuming a hierarchical ordering of the endogenous variables in which the variable ordered first is not contemporaneously affected by the others. The variable ordered second is assumed to be contemporaneously affected only by the variable ordered first and so on.

factor VAR' - includes the first four principal components (for the ask and bid side) along with the mid-quote change as endogenous variables.

For each stock the reduced form VARs are estimated equation by equation by OLS. Information criteria suggest a VAR lag order equal to two. The variance decomposition focuses on the 15-step-ahead forecast error of the mid-quote change.[13] Consistent with similar market microstructure papers that use time series models (Chan (2005); Giot and Grammig (2006)), previous day information is not allowed to affect the present day. Observations involving overnight mid-quote changes are excluded. Furthermore, if the lags of the VAR equations reach back to the previous day, then these lagged values are substituted by sample means.

4.6.1. Visible factor VAR

Table 9 contains the results for the 'visible factor VAR'. The table reports the information shares (averaged over the thirty stocks) and the corresponding cross-sectional standard deviations. The results show that the shift and rotation factors have informational content with respect to the long-run movements of the mid-quote. Averaged over the four Cholesky orderings, the aggregated information share of the principal components amounts to 7.5%. It is not surprising that the own information share of the mid-quote innovations dominates. The innovations in the structural equation for the mid-quote change account for new public and private information brought about by trades occurring within the five-minute sampling interval as well as non-informationally related microstructure effects like inventory accounting and rounding errors.[14] Beyond the 'own information share' effect, we can nevertheless conclude that the shape of the order book contains additional informational content with respect to the long-run movements of the asset price.[15]

When the information shares of the first principal components are compared, the results in Table 9 suggest that the shift factor is more informative than the rotation factor. However, the cross-sectional standard deviations of the variance shares are quite large. This indicates that studying on a finer aggregation level may be warranted. For that purpose, Table 10

[13]The convergence of the variance of the forecast error to its unconditional variance is actually faster. For most stocks decomposing the 5-step-ahead forecast yields virtually the same results.

[14]These microstructure effects should not matter to a large extent. In Xetra there are no dedicated market makers who have to carry involuntary inventory. Furthermore the minimum tick size is just one euro cent.

[15]In a related paper for the Australian Stock Exchange, Cao, Hansch, and Wang (2004) argue that the information share of the book beyond the best quotes amounts to 30%. However, their econometric methodology is different as they estimate a cointegrated system for the share prices (and not relative price impacts).

reports cross-sectional averages and standard deviations for stocks grouped by trading activity. Specifically, we sort the 30 stocks into four trading activity quartiles. Trading activity is measured in average number of trades per day. The results show a more differentiated picture. First, the information share of the principal components is higher for less actively traded stocks. The total variance share of the factors (averaged across orderings) amounts to 9.9% for the group of least frequently traded stocks, while for the actively traded stocks (first quartile) the average variance share is 5.3%. The aggregated information shares for the second quartile (average 6.4%) and third quartile (average 7.4%) are in between. This suggests that the informational content of the book is higher for less actively traded stocks. This is consistent with recent evidence documented by Naes and Skjeltorp (2005), who focus on the trading volume - volatility relationship in order book markets. They show that this relationship is shaped by the fact there is more disagreement regarding the future earnings of small-cap stocks (than large-cap stocks). As a result, order books tend to be more informative for less actively stocks than for large stocks.

In the same vein, the proportion of information shares attributable to the shifts and rotations of the book is different for actively and less actively traded stocks. Table 10 shows that, for the group of actively traded stocks, the shift factor accounts for 88% of the total information share attributable to the principal components (sum of ask and bid side information shares averaged across the orderings). However, for the least active quartile, rotations of the order book are relatively more informative: the rotation factor accounts for 57% of the aggregated information share attributable to the principal components (again sum of ask and bid side shares averaged across orderings). The second activity quartile is similar to the most active quartile with the shift factor accounting for 85% of the total information share. For the third activity quartile, the information shares are more balanced (55% for the shift factor, 45% for the rotation factor). The result that the slope of the order book (determined by the rotation factor) is more informative for less actively traded stocks again corroborates the findings in Naes and Skjeltorp (2005).

4.6.2. Visible and hidden factor VAR results

In order to assess the informational content of the hidden orders, we now turn to the 'visible and hidden factor VAR'. In this case, we use the first four ask and bid side principal components obtained from PCAs performed on visible price impacts and the incremental price impacts brought about by hidden orders. As discussed in the previous section, the first two of the four factors assume the role of shifting the price impact curve, while the third and fourth factors account for the change of the slope of the price impact curve. Table 11 confirms the results reported for the 'visible factors VAR' in that the total information share of the principal components is higher for less actively traded stocks. Furthermore, as discussed above, we find that for actively traded stocks shifts of the order book seem more informative. For less frequently traded stocks, a rotation of the book slope carries more informational content. However, the information share of all (eight) principal components combined is not substantially increased compared to the 'visible factors' specification. Averaged across orderings, the aggregated variance share attributable to the principal components increases from 5.3% to 6.3% for the most active quartile and from 9.9% to 10.7% for the least actively traded quartile. For the other two quartiles the results are similar (from 6.4% to 7.6% for the second quartile and from 7.4% to 8.3% for the third quartile). This suggests that hidden orders contain some informational content in excess of shifts and rotations of the visible book, but that the majority of the informational content is already accounted for by the visible factors. This is in agreement with De Winne and D'Hondt (2005). Using a different empirical methodology and Euronext order book data, they conclude that hidden orders are not related to informed trading.

5. Conclusion

This paper was motivated by theoretical models which suggest that a small set of latent factors accounts for the variation of liquidity in an order book. Put in statistical terms, microstructure theory hints at the existence of order book commonalities. The recent availability of order book data called for an empirical investigation of this issue. To conduct such an analysis, we have used high quality order book data from one of the largest European stock markets, more precisely reconstructed limit order book sequences for the thirty German blue chip stocks that constitute the DAX30 index. By performing principal components analysis on the reconstructed order books we investigated the heterogeneity liquidity suppliers. The stock specific analysis was augmented by a more traditional cross sectional analysis of liquidity commonalities. Finally, we have also quantified the informational content of the extracted factors, as well as that of the hidden liquidity. The following findings are of particular importance:

First, two principal components already provide a considerable explanatory power to account for the inter-temporal variation of stock specific order books. This holds true for 'one-sided' analysis, i.e. when the principal component analysis is performed separately for the buy and the sell side of the order book. We showed that the two extracted factors have clearly identified purposes: Variations in the first principal component account for nonlinear shifts of the liquidity supply curves, while the second principal component rotates the slope of the price impact curves.

Second, the paper showed that bid and ask side of the book on the one hand, and visible and hidden portions of liquidity on the other, are driven by distinct, albeit correlated factors. We interpret this as evidence that an open order book market attracts a heterogeneous population of limit order traders, with heterogeneity referring to asset valuation, impatience and being informed about the true asset value.

Third, and in line with previous findings, we also found evidence of liquidity commonality across stocks. However, the total explanatory power of the principal components is considerably smaller than in the stock specific analysis. Our cross sectional commonality results are, however, broadly comparable to those reported in Hasbrouck and Seppi (2001) and Bauer (2004).

Fourth, we showed by means of a VAR analysis that the information share attributable to the extracted factors with respect to the long run evolution of the asset price is non-negligible. The information shares have a considerably cross sectional variation, but a clear pattern is discernible. While for the group of least frequently traded stocks, we estimated an information share of the principal components of about ten percent, the information share for the group of most actively traded stocks is only half the size.

Fifth, extending the VAR analysis to account for the hidden liquidity factors we found that the hidden part of the book does not have economically significant informational content.

Avenues for further research stretch in various directions. The heterogeneity of limit order traders is clearly an issue that warrants further investigation. By means of a bivariate GARCH model one could investigate the time series properties of cross-side correlations of price impacts and their determinants. The small-big stock analysis also calls for further research. What is the economic story behind the result that for less frequently traded stocks changes in the slope of the book seem more informative, while for actively traded stocks it is the shift of the price impact curves that conveys informational content? Furthermore, an analysis along the lines of Gomber, Schweickert, and Theissen (2004) could investigate the effect of volatility and liquidity demand shocks (large trades hitting one side of the market) on the shift and rotation factor, thereby providing a description how the order book shape responds towards liquidity demand shocks. Finally, the VAR analysis conducted in this paper could be extended to study the dynamics of various dimensions of liquidity supply. Besides the shift and rotation factors one could include the inside spread as an endogenous variable and study the dynamics of inside spread and beyond best quote book liquidity.

References

Ahn, H., K. Bae, and K. Chan (2001): "Limit orders, depth and volatility: evidence from the stock exchange of Hong Kong," *Journal of Finance*, 56, 767–788.

Ahn, H., and Y. Cheung (1999): "The intraday patterns of the spread and depth in a market with market makers: the Stock Exchange of Hong Kong," *Pacific-Basin Finance Journal*, 7, 539–556.

Bauer, W. (2004): "Commonality in liquidity in pure order-driven markets," Mimeo.

Biais, B., P. Hillion, and C. Spatt (1995): "An empirical analysis of the limit order book and the order flow in the Paris Bourse," *Journal of Finance*, 50, 1655–1689.

Brockman, P., and D. Chung (1999): "An analysis of depth behavior in an electronic, order-driven environment," *Journal of Banking and Finance*, 23, 1861–1886.

——— (2002): "Commonality in liquidity: evidence from an order-driven market structure," *Journal of Financial Research*, 25, 521–539.

Cao, C., O. Hansch, and X. Wang (2004): "The informational content of an open limit order book," AFA 2005 Philadelphia Meetings; EFA 2004 Maastricht Meetings Paper No. 4311.

Chan, Y. (2005): "Price movement effects on the state of the electronic limit-order book," *The Financial Review*, 40, 195–221.

Chordia, T., R. Roll, and A. Subrahmanyam (2001): "Commonality in liquidity," *Journal of Financial Economics*, 56, 3–28.

Coughenour, J., and M. Saad (2004): "Common market makers and commonality in liquidity," *Journal of Financial Economics*, 73, 37–69.

Danielsson, J., and R. Payne (2001): "Measuring and explaining liquidity on an electronic limit order book: evidence from Reuters D2000-2," Financial Markets Group, London School of Economics.

De Winne, R., and C. D'Hondt (2005): "Market transparency and traders' behavior: An analysis on Euronext with full order book data," Working paper, Catholic Unversity of Mons and EDHEC Business School, Lille.

Deutsche Börse AG (1999): "Xetra Market Model Release 3 Stock Trading," Technical Report.

Domowitz, I., O. Hansch, and X. Wang (2005): "Liquidity commonality and return co-movement," *Forthcoming in Journal of Financial Markets*.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2003): "The generalized factor model: consistency and rates," *Forthcoming in the Journal of Econometrics*.

Foucault, T., O. Kadan, and E. Kandel (2003): "Limit order book as a market for liquidity," *Forthcoming, Review of Financial Studies*.

Giot, P., and J. Grammig (2006): "How large is liquidity risk in an automated auction market?," *Empirical Economics*, 30(4), 867 – 887.

Glosten, L. (1994): "Is the electronic open limit order book inevitable?," *Journal of Finance*, 49, 1127–1161.

Gomber, P., U. Schweickert, and E. Theissen (2004): "Zooming in on Liquidity," EFA 2004 Working Paper, University of Bonn.

HAMILTON, J. D. (1994): *Time Series Analysis.* Princeton University Press.

HANSCH, O. (2003): "Island tides: exploring ECN liquidity," Mimeo.

HASBROUCK, J. (1991): "Measuring the information content of stock trades," *Journal of Finance*, 46(1), 179–207.

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

IRVINE, P., G. BENSTON, AND E. KANDEL (2000): "Liquidity beyond the inside spread: measuring and using information in the limit order book," Mimeo, Goizueta Business School, Emory University, Atlanta.

KEMPF, A., AND D. MAYSTON (2006): "Systematic Liquidity across Limit Order Books - Evidence from Order Data," Mimio. University of Cologne.

NAES, R., AND J. SKJELTORP (2005): "Order book characteristics and the volume-volatility relation: Empirical evidence from a limit order market," Working Paper Norges Bank and Norwegian School of Managment.

PARDO, A., AND R. PASCUAL (2004): "On the hidden side of liquidity," Working Paper Universidad de las Islas Baleares.

PARLOUR, C. A. (1998): "Price Dynamics in Limit Order Markets," *Review of Financial Studies*, 11, 789–816.

STOLL, H. (1978): "The supply of dealer services in securities markets," *Journal of Finance*, 33, 1133–1151.

Table 1: **Characteristics of the DAX30 stocks.**

Ticker	Company Name	Daily Turnover	Market cap (mill.)	Trade size	% agg trades	Daily nb trades	Daily nb subm.	Daily nb cancel.	Price	Spread	Spread %
DTE	DEUTSCHE TELEKOM	350627866	34858	78884	5.0	4445	14498	11009	15.7	0.01	0.07
SIE	SIEMENS	321704299	52893	72831	16.7	4418	23659	19920	64.0	0.03	0.05
DBK	DEUTSCHE BANK	309282831	38228	78083	19.3	3961	23169	19772	67.2	0.03	0.05
ALV	ALLIANZ	289980556	33805	64114	21.4	4523	29791	25882	100.1	0.05	0.05
MUV2	MUENCHENER RUECK	207353230	16396	60534	20.7	3425	20154	16894	93.9	0.06	0.06
DCX	DAIMLER CHRYSLER	187737846	30316	56736	14.5	3309	18722	15919	36.4	0.02	0.06
SAP	SAP	184628162	27412	65795	21.9	2806	19733	17095	131.5	0.08	0.06
EOA	E.ON	160625983	33753	55950	13.6	2871	18899	16468	52.5	0.03	0.06
IFX	INFINEON	146462315	4790	52331	8.6	2799	10320	7744	11.6	0.01	0.10
BAS	BASF	124434537	25425	48236	13.8	2580	18211	15898	43.3	0.03	0.06
VOW	VOLKSWAGEN	104249843	9688	40963	16.0	2545	13474	11273	39.2	0.03	0.07
HVM	BAY.HYPO-VEREINSBK.	98351090	6629	50783	15.0	1937	10204	8293	18.7	0.02	0.11
RWE	RWE	97655566	12653	42203	13.0	2314	14438	12355	33.8	0.03	0.08
BAY	BAYER	88776121	15911	36994	12.4	2400	15258	12988	23.1	0.02	0.08
BMW	BMW	87854358	12211	41639	14.4	2110	14736	12764	34.7	0.02	0.07
CBK	COMMERZBANK	53171668	7569	36659	12.6	1450	11922	10476	15.4	0.02	0.11
SCH	SCHERING	51413053	7055	33756	16.2	1523	9111	7669	40.8	0.04	0.09
LHA	LUFTHANSA	43946809	4548	32504	11.9	1352	8079	6780	14.2	0.02	0.12
DPW	DEUTSCHE POST	43836617	6806	33330	11.0	1315	6861	5666	18.2	0.02	0.11
MEO	METRO	38874669	5018	31480	15.7	1235	7975	6702	35.0	0.04	0.12
TKA	THYSSEN KRUPP	37892493	6450	30017	11.3	1262	7864	6672	15.9	0.02	0.13
DB1	DEUTSCHE BOERSE	35696903	4847	36359	18.4	982	6598	5698	46.9	0.04	0.10
ADS	ADIDAS-SALOMON	31976047	4104	32635	20.1	980	8057	7105	92.6	0.08	0.09
ALT	ALTANA	30985416	3338	28310	18.9	1095	7718	6609	48.6	0.05	0.10
MAN	MAN	27685031	2434	26189	13.0	1057	7214	6235	27.7	0.03	0.12
TUI	TUI	26281175	2025	24723	17.6	1063	6767	5714	18.7	0.03	0.14
CONT	CONTINENTAL	25627638	4060	25574	13.5	1002	8036	7052	31.6	0.04	0.11
LIN	LINDE	22378772	3448	24971	15.8	896	8342	7454	43.6	0.05	0.11
HEN3	HENKEL	18174548	3682	25904	16.6	702	7989	7306	65.9	0.07	0.10
FME	FRESENIUS MEDICAL CARE	12850947	1944	20680	16.7	621	5764	5195	54.0	0.07	0.13
	Average	108683880	14076	42972	15.2	2099	12785	10887	44.5	0.04	0.09

Characteristics of the stocks included in the DAX30 index. For each stock, *Market cap* is the market capitalization in million euros at the end of December 2003, *Trade size* is the average trade size over the 3-month sample, and *%agg trades* is the percentage of total trading volume that walked up the book. The remaining statistics are daily averages over the 3 months: *Daily turnover* is the total amount of shares traded in euros per trading day, *Daily nb trades* is the daily number of trades, *Daily nb subm.* is the daily number of order submissions, market orders excluded, and *Daily nb cancel.* is the daily number of order cancellations. Finally, *Price*, *Spread* and *Spread, %* refer to the average mid-quote, spread and relative spread over the 3 months. The sample ranges from January 2, 2003 to March 31, 2003.

Table 2: Descriptive statistics for the trade sizes

quantile	mean	std. deviation	min	max
0.025	996	271	552	1577
0.05	1846	524	1042	3127
0.075	2698	722	1562	4396
0.1	3478	1026	1880	6296
0.125	4111	1163	2012	6896
0.15	4862	1428	2742	8312
0.175	5830	2005	3388	10666
0.2	6704	2486	3698	12851
0.225	7567	2682	3960	13906
0.25	8582	3126	4713	16217
0.275	9755	3880	5234	19125
0.3	10973	4553	5470	20895
0.325	12274	5150	5578	23991
0.35	13522	5679	5886	26424
0.375	14839	6139	6606	28593
0.4	16335	6652	7579	31595
0.425	17828	7021	8520	33260
0.45	19478	7535	9566	34835
0.475	21145	8186	10300	37446
0.5	22857	9130	10837	40854
0.525	24867	10031	11266	45356
0.55	27021	10908	12622	50365
0.575	29552	11923	14184	55451
0.9	97845	36988	50346	176198
0.925	116841	47979	54670	235050
0.95	146515	64124	63703	316026
0.975	203758	87956	91440	441615
0.99	310684	147832	133775	733206
1	3350844	2659970	1016500	13578391

Descriptive statistics for the quantiles of the trade sizes expressed in euros. For the quantile given at the start of the row, the next columns give the mean, standard deviation, minimum and maximum trade sizes in euros (on average for the 30 stocks).

Table 3: Descriptive statistics for the price impacts

	mean	std. deviation	dispersion index
20,000	0.01	0.02	2.81
50,000	0.02	0.03	1.82
100,000	0.03	0.04	1.36
500,000	0.18	0.16	0.87

Descriptive statistics for the price impacts. Each row corresponds to a complete book ask price impact series, for a hypothetical trade size of 20,000, 50,000, 100,000 and 500,000 euros. The columns read as follows: *mean* is the average across the 30 stocks of the mean of the price impact, *st.deviation* is the average across the 30 stocks of the standard deviation of the price impact, and *dispersion index* is the average across the 30 stocks of the dispersion index of the price impact. The dispersion index is the ratio of the standard deviation to the mean of the series.

Table 4: Correlations between price impact series

variables	avg	std	min	max
PIA 3 — PIA 11	0.66	0.03	0.59	0.72
PIA 3 — PIA 51	0.31	0.05	0.20	0.38
PIA 3 — PIB 3	0.01	0.02	-0.02	0.05
PIA 3 — PIB 11	0.03	0.03	-0.04	0.08
PIA 3 — PIB 51	0.05	0.04	-0.06	0.14
DIFF1 PIA 3 — PIA 11	-0.41	0.05	-0.48	-0.28
DIFF1 PIA 3 — PIA 51	-0.15	0.05	-0.23	-0.07
DIFF1 PIA 3 — PIB 3	0.01	0.01	-0.01	0.05
DIFF1 PIA 3 — PIB 11	0.01	0.02	-0.02	0.05
DIFF1 PIA 3 — PIB 51	0.00	0.01	-0.02	0.04
PIA 11 — PIA 3	0.66	0.03	0.59	0.72
PIA 11 — PIA 51	0.62	0.05	0.48	0.68
PIA 11 — PIB 3	0.03	0.04	-0.03	0.12
PIA 11 — PIB 11	0.09	0.06	-0.07	0.22
PIA 11 — PIB 51	0.14	0.08	-0.08	0.35
DIFF1 PIA 11 — PIA 3	-0.45	0.03	-0.50	-0.38
DIFF1 PIA 11 — PIA 51	-0.29	0.08	-0.43	-0.16
DIFF1 PIA 11 — PIB 3	0.01	0.01	-0.02	0.03
DIFF1 PIA 11 — PIB 11	0.00	0.02	-0.03	0.06
DIFF1 PIA 11 — PIB 51	0.00	0.02	-0.03	0.08
PIA 51 — PIA 3	0.31	0.05	0.20	0.38
PIA 51 — PIA 11	0.62	0.05	0.48	0.68
PIA 51 — PIB 3	0.06	0.04	-0.04	0.12
PIA 51 — PIB 11	0.14	0.07	-0.07	0.29
PIA 51 — PIB 51	0.22	0.10	-0.03	0.62
DIFF1 PIA 51 — PIA 3	-0.22	0.04	-0.27	-0.13
DIFF1 PIA 51 — PIA 11	-0.39	0.06	-0.47	-0.24
DIFF1 PIA 51 — PIB 3	0.01	0.01	-0.02	0.04
DIFF1 PIA 51 — PIB 11	0.00	0.02	-0.04	0.07
DIFF1 PIA 51 — PIB 51	-0.01	0.02	-0.04	0.09

Correlations between the complete price impact series at the bid and ask side. PIA stands for ask price impacts, PIB stands for bid price impacts. PIA 3 corresponds to an ask price impact for a trade size of 20,000 euros, PIA 11 100,000 euros, and PIA 51 500,000 euros (the same holds for the bid price impacts PIB 3, PIB 11, PIB 51). *DIFF1* means that the variable is differentiated as e.g. $DIFF1PIA3 = PIA3_{t+1} - PIA3_t$. Variables not differentiated are observed at time t. *Avg* is the average correlation across the 30 stocks, *Std* its standard deviation, and *Min* and *Max* the smallest and largest correlations across the 30 stocks. For example, an entry reads as follows: the correlation between the ask price impact for a trade size of 20,000 euros (PIA 3) and the ask price impact for a trade of 100,000 euros (PIA 11) is on average equal to 0.66.

Table 5: Spread correlation

variable	correlation
PIA 20,000	-0.04
PIA 100,000	0.02
PIA 500,000	0.13
PIB 20,000	-0.05
PIB 100,000	0.01
PIB 500,000	0.11
XLM A 20,000	0.86
XLM A 100,000	0.62
XLM A 500,000	0.37
XLM B 20,000	0.85
XLM B 100,000	0.61
XLM B 500,000	0.35
XLM 20,000	0.92
XLM 100,000	0.72
XLM 500,000	0.44

Five-minute contemporaneous correlation between the spread and the variable given in the first column (on average for the 30 stocks). PIA stands for ask complete price impacts, PIB stands for bid complete price impacts. XLM A and XLM B are the ask and bid price impacts normalized by the mid-quote, while XLM is the sum of XLM A and XLM B (i.e. XLM is the volume-weighted spread). The number next to PIA, PIB, XLM A, XLM B or XLM gives the trade size in euros.

Table 6: Principal component analysis for the price impacts

	Visible price impacts								
	Ask					Bid			
	Avg	Stdv	Min	Max		Avg	Stdv	Min	Max
$F1$	0.82	0.02	0.74	0.85	$F1$	0.82	0.03	0.74	0.85
$F2$	0.12	0.02	0.10	0.17	$F2$	0.13	0.02	0.10	0.18
$F3$	0.03	0.01	0.03	0.05	$F3$	0.03	0.00	0.03	0.05
$F4$	0.01	0.00	0.01	0.02	$F4$	0.01	0.00	0.01	0.02
	Complete price impacts								
	Ask					Bid			
	Avg	Stdv	Min	Max		Avg	Stdv	Min	Max
$F1$	0.82	0.02	0.75	0.85	$F1$	0.82	0.03	0.74	0.85
$F2$	0.12	0.01	0.10	0.16	$F2$	0.12	0.02	0.10	0.18
$F3$	0.03	0.00	0.03	0.05	$F3$	0.03	0.00	0.03	0.05
$F4$	0.01	0.00	0.01	0.02	$F4$	0.01	0.00	0.01	0.02
	Visible and hidden price impacts								
	Ask					Bid			
	Avg	Stdv	Min	Max		Avg	Stdv	Min	Max
$F1$	0.50	0.03	0.42	0.57	$F1$	0.49	0.04	0.40	0.55
$F2$	0.30	0.03	0.25	0.36	$F2$	0.31	0.03	0.26	0.36
$F3$	0.09	0.01	0.08	0.11	F3	0.08	0.01	0.07	0.10
$F4$	0.05	0.01	0.04	0.07	$F4$	0.05	0.01	0.04	0.08
	Bid and ask price impacts								
	Visible					Complete			
	Avg	Stdv	Min	Max		Avg	Stdv	Min	Max
$F1$	0.47	0.02	0.43	0.50	$F1$	0.46	0.02	0.43	0.50
$F2$	0.35	0.03	0.29	0.41	$F2$	0.36	0.03	0.31	0.41
$F3$	0.06	0.01	0.06	0.09	$F3$	0.06	0.01	0.05	0.09
F4	0.06	0.01	0.05	0.08	$F4$	0.06	0.01	0.05	0.08

Explanatory power for the first 4 principal components in the PCA. The first panel (starting from the top) reports the explanatory power of a PCA on the visible ask price impacts (left side) and of a PCA on the visible bid price impacts (right side). $F1$ to $F4$ refer to the explanatory power for the first to fourth principal components respectively. The second panel displays similar results except that the PCA is applied to the complete price impacts. The third panel reports the explanatory powers for a PCA on the visible and hidden price impacts, and the bottom panel for a PCA on both sides of the book. *Avg* is the average explanatory power across the 30 stocks, *Std* its standard deviation, and *Min* and *Max* the smallest and largest explanatory power across the 30 stocks.

Table 7: Principal component analysis for the price impact differences

	Visible price impacts							
	Ask				Bid			
	Avg	Stdv	Min	Max	Avg	Stdv	Min	Max
$F1$	0.52	0.04	0.38	0.58	0.51	0.04	0.38	0.58
$F2$	0.15	0.01	0.14	0.19	0.16	0.01	0.13	0.19
$F3$	0.08	0.01	0.07	0.11	0.09	0.01	0.07	0.11
$F4$	0.05	0.00	0.05	0.07	0.05	0.00	0.05	0.07
	Complete price impacts							
	Ask				Bid			
	Avg	Stdv	Min	Max	Avg	Stdv	Min	Max
$F1$	0.54	0.04	0.39	0.59	0.53	0.04	0.39	0.60
$F2$	0.16	0.01	0.13	0.19	0.16	0.01	0.12	0.19
$F3$	0.08	0.01	0.07	0.11	0.08	0.01	0.07	0.11
$F4$	0.05	0.00	0.04	0.07	0.05	0.00	0.04	0.07
	Visible and hidden price impacts							
	Ask				Bid			
	Avg	Stdv	Min	Max	Avg	Stdv	Min	Max
$F1$	0.34	0.03	0.24	0.39	0.33	0.03	0.23	0.40
$F2$	0.18	0.03	0.13	0.24	0.19	0.03	0.14	0.24
$F3$	0.11	0.01	0.08	0.13	0.11	0.01	0.08	0.13
$F4$	0.06	0.00	0.05	0.07	0.06	0.01	0.05	0.08
	Bid and ask price impacts							
	Visible				Complete			
	Avg	Stdv	Min	Max	Avg	Stdv	Min	Max
$F1$	0.31	0.03	0.20	0.40	0.31	0.03	0.21	0.40
$F2$	0.21	0.02	0.18	0.25	0.23	0.02	0.18	0.26
$F3$	0.08	0.01	0.07	0.10	0.08	0.01	0.07	0.10
$F4$	0.08	0.01	0.06	0.09	0.08	0.01	0.06	0.09

Explanatory power for the first 4 principal components in the PCA. The first panel (starting from the top) reports the explanatory power of a PCA on the visible ask price impact differences (left side) and of a PCA on the visible bid price impact differences (right side). $F1$ to $F4$ refer to the explanatory power for the first to fourth principal components respectively. The second panel displays similar results except that the PCA is applied to the complete price impact differences. The third panel reports the explanatory powers for a PCA on the visible and hidden price impact differences, and the bottom panel for a PCA on both sides of the book. *Avg* is the average explanatory power across the 30 stocks, *Std* its standard deviation, and *Min* and *Max* the smallest and largest explanatory power across the 30 stocks.

Table 8: Cross-sectional principal component analysis

	Factors identified in Section 4.1				
	Ask			Bid	
	FA1	FA2		FB1	FB2
$F1$	0.10	0.05	$F1$	0.08	0.05
$F2$	0.04	0.04	$F2$	0.05	0.04
$F3$	0.04	0.04	$F3$	0.04	0.04
$F4$	0.04	0.04	$F4$	0.04	0.04
	Order book liquidity variables				
	10,000-euro PI	100,000-euro PI		Spread	Depth
$F1$	0.04	0.07	$F1$	0.09	0.03
$F2$	0.04	0.04	$F2$	0.04	0.02
$F3$	0.04	0.04	$F3$	0.04	0.02
$F4$	0.04	0.04	$F4$	0.04	0.02

Explanatory power for the first 4 principal components in the PCA (cross-sectional analysis on the 30 stocks). The top panel reports the explanatory power of a cross-sectional PCA on the first (FA1 for the ask side and FB1 for the bid side) and second factors (FA2 and FB2) identified in Section 4.1. $F1$ to $F4$ refer to the explanatory power for the first to fourth principal components respectively. The bottom panel reports the explanatory power of a cross-sectional PCA on some liquidity variables. *10,000-euro PI* (*100,000-euro PI*) refers to the complete ask price impacts for a trade size of 10,000 (100,000) euros, *Spread* to the usual bid-ask spread in basis points, and *Depth* to the depth displayed at the best ask price.

Cholesky Ordering	$F1A$	$F1B$	$F2A$	$F2B$	Σ	own
$F1 - F2 - MQ$	3.9	3.0	2.0	2.1	11.0	89.0
	(2.9)	(1.2)	(2.1)	(2.4)	(8.5)	(3.9)
$F2 - F1 - MQ$	3.6	2.8	2.3	2.3	11.0	89.0
	(2.9)	(1.6)	(1.9)	(2.1)	(8.5)	(3.9)
$MQ - F1 - F2$	1.3	1.6	0.3	0.3	3.5	96.5
	(0.7)	(0.7)	(0.4)	(0.4)	(2.2)	(1.6)
$MQ - F2 - F1$	0.8	0.9	0.8	1.0	3.5	96.5
	(0.4)	(0.5)	(0.7)	(0.9)	(2.5)	(1.6)

Table 9: **Information shares of visible factors with respect to mid-quote movements (averages across all stocks)** We estimate a structural VAR that includes the mid-quote change and the first two ask and bid factors as endogenous variables. The factors are obtained from separate PCAs of the ask and bid side visible price impacts. The sampling frequency is five minutes. The table reports the results of a variance decomposition of the long-run forecasting error of the mid-quote change. The information shares (in %) attributable to the first (shift) and second (rotation) factor factor are reported in the columns $F1A$, $F2A$ (ask side) and $F1B$ and $F2B$ (bid side). The column Σ reports the aggregated information share of all factors. The information share of the mid-quote innovations is reported in column *own*. The variance shares are averaged across all stocks. The numbers in parentheses are cross-sectional standard deviations. The table shows the results for four alternative orderings for a Cholesky decomposition of the variance covariance matrix of the residuals of the VAR in standard form. The orderings interchange the position of the mid-quote change (MQ) and the first and second factors. Within each ordering the ask factors are placed before the bid side factors. For example the ordering $F1 - F2 - MQ$ (shift factors before rotation factors before mid-quote change) reads in detail $F1A - F1B - F2A - F2B - MQ$.

	Cholesky Ordering	$F1A$	$F1B$	$F2A$	$F2B$	Σ	own
1^{st} quartile (most active)	$F1-F2-MQ$	4.2	3.1	1.0	0.8	9.1	90.9
		(1.2)	(1.0)	(0.5)	(0.5)	(3.2)	(2.8)
	$F2-F1-MQ$	4.8	3.6	0.4	0.3	9.1	90.9
		(1.5)	(1.3)	(0.2)	(0.2)	(3.3)	(2.8)
	$MQ-F1-F2$	0.6	0.8	0.0	0.0	1.5	98.5
		(0.3)	(0.3)	(0.0)	(0.0)	(0.7)	(0.5)
	$MQ-F2-F1$	0.6	0.8	0.1	0.1	1.5	98.5
		(0.3)	(0.3)	(0.1)	(0.1)	(0.8)	(0.5)
2^{nd} quartile	$F1-F2-MQ$	5.3	3.6	0.5	0.4	9.8	90.2
		(3.8)	(1.7)	(0.6)	(0.5)	(6.5)	(5.0)
	$F2-F1-MQ$	4.9	3.0	1.0	1.0	9.8	90.2
		(4.2)	(1.8)	(0.7)	(0.6)	(7.3)	(5.0)
	$MQ-F1-F2$	1.3	1.6	0.1	0.1	3.0	97.0
		(0.5)	(0.6)	(0.1)	(0.0)	(1.2)	(0.9)
	$MQ-F2-F1$	1.0	1.2	0.4	0.5	3.0	97.0
		(0.4)	(0.5)	(0.3)	(0.2)	(1.4)	(0.9)
3^{rd} quartile	$F1-F2-MQ$	3.8	2.6	2.3	2.1	10.8	89.2
		(3.4)	(1.0)	(1.5)	(1.7)	(7.6)	(3.1)
	$F2-F1-MQ$	2.5	1.5	3.8	3.0	10.8	89.2
		(2.8)	(0.8)	(1.1)	(1.1)	(5.9)	(3.1)
	$MQ-F1-F2$	1.6	1.9	0.3	0.4	4.1	95.9
		(0.2)	(0.4)	(0.1)	(0.3)	(1.0)	(0.7)
	$MQ-F2-F1$	0.8	0.9	1.1	1.2	4.1	95.9
		(0.2)	(0.4)	(0.3)	(0.3)	(1.2)	(0.7)
4^{th} quartile (least active)	$F1-F2-MQ$	2.2	2.9	4.2	5.3	14.6	85.4
		(1.5)	(0.7)	(2.7)	(2.5)	(7.3)	(2.2)
	$F2-F1-MQ$	2.1	3.4	4.2	4.9	14.6	85.4
		(1.3)	(1.5)	(1.6)	(1.8)	(6.2)	(2.2)
	$MQ-F1-F2$	1.6	2.2	0.8	0.8	5.3	94.7
		(1.0)	(0.8)	(0.7)	(0.4)	(2.9)	(1.4)
	$MQ-F2-F1$	0.7	0.8	1.7	2.1	5.3	94.7
		(0.5)	(0.7)	(0.7)	(0.8)	(2.7)	(1.4)

Table 10: **Information shares of visible factors with respect to mid-quote movements (trading activity quartiles)** We estimate a structural VAR that includes the mid-quote change and the first two ask and bid factors as endogenous variables. The factors are obtained from separate PCAs of the ask and bid side visible price impacts. The table reports the results of a variance decomposition of the long-run forecasting error of the mid-quote change. The information shares (in %) attributable to the first (shift) and second (rotation) factor factor are reported in the columns $F1A$, $F2A$ (ask side) and $F1B$ and $F2B$ (bid side). The column Σ reports the aggregated information share of all factors. The information share of the mid-quote innovations is found in column *own*. The variance shares are averaged across all stocks within the same trading activity (measured in trades per day) quartile. The numbers in parentheses are cross-sectional standard deviations. The table shows the results for four alternative orderings for a Cholesky decomposition of the variance covariance matrix of the residuals of the VAR in standard form. The orderings interchange the position of the mid-quote change (MQ) and the first and second factors. Within each ordering the ask factors are placed before the bid side factors. For example, the ordering $F1-F2-MQ$ (shift factors before rotation factors before mid-quote change) reads in detail $F1A-F1B-F2A-F2B-MQ$.

Quartile	Cholesky Ordering	$F1$	$F2$	$F3$	$F4$	Σ	MQ
1^{st} quartile (most active)	$F1-F2-F3-F4-MQ$	3.3	4.8	0.9	1.1	10.1	89.9
		(1.8)	(1.1)	(0.6)	(0.7)	(4.1)	(2.5)
	$F3-F4-F1-F2-MQ$	3.9	5.3	0.5	0.4	10.1	89.9
		(2.2)	(1.1)	(0.3)	(0.4)	(4.0)	(2.5)
	$MQ-F1-F2-F3-F4$	0.3	1.9	0.1	0.2	2.4	97.6
		(0.2)	(0.6)	(0.1)	(0.1)	(1.0)	(0.4)
	$MQ-F3-F4-F1-F2$	0.2	1.9	0.2	0.1	2.4	97.6
		(0.2)	(0.6)	(0.2)	(0.1)	(1.0)	(0.4)
2^{nd} quartile	$F1-F2-F3-F4-MQ$	4.1	6.0	0.6	0.4	11.1	88.9
		(3.5)	(2.7)	(0.7)	(0.3)	(7.2)	(4.6)
	$F3-F4-F1-F2-MQ$	3.7	5.7	0.8	0.9	11.1	88.9
		(3.7)	(2.9)	(0.6)	(0.7)	(7.8)	(4.6)
	$MQ-F1-F2-F3-F4$	0.6	3.4	0.1	0.1	4.2	95.8
		(0.2)	(1.2)	(0.1)	(0.1)	(1.6)	(1.1)
	$MQ-F3-F4-F1-F2$	0.4	3.1	0.3	0.4	4.2	95.8
		(0.2)	(1.1)	(0.2)	(0.3)	(1.8)	(1.1)
3^{rd} quartile	$F1-F2-F3-F4-MQ$	2.2	5.2	2.0	2.1	11.5	88.5
		(2.0)	(2.9)	(1.8)	(1.8)	(8.4)	(3.4)
	$F3-F4-F1-F2-MQ$	1.7	3.8	2.7	3.3	11.5	88.5
		(1.8)	(2.8)	(2.0)	(1.3)	(7.9)	(3.4)
	$MQ-F1-F2-F3-F4$	0.8	3.6	0.2	0.5	5.1	94.9
		(0.4)	(1.0)	(0.2)	(0.3)	(2.0)	(0.7)
	$MQ-F3-F4-F1-F2$	0.4	2.4	0.7	1.6	5.1	94.9
		(0.3)	(1.1)	(0.4)	(0.4)	(2.2)	(0.7)
4^{th} quartile (least active)	$F1-F2-F3-F4-MQ$	2.6	3.9	3.2	5.6	15.4	84.6
		(1.1)	(2.6)	(2.1)	(3.9)	(9.6)	(1.4)
	$F3-F4-F1-F2-MQ$	3.7	3.1	2.5	6.1	15.4	84.6
		(2.2)	(2.0)	(2.1)	(3.3)	(9.6)	(1.4)
	$MQ-F1-F2-F3-F4$	1.4	2.9	0.5	1.1	6.0	94.0
		(0.7)	(1.5)	(0.8)	(1.0)	(4.0)	(1.3)
	$MQ-F3-F4-F1-F2$	0.6	1.7	1.0	2.7	6.0	94.0
		(0.5)	(1.5)	(1.0)	(1.8)	(4.7)	(1.3)

Table 11: **Incremental informational content of hidden orders** We estimate a structural VAR that includes the mid-quote change and four ask and bid factors. These factors are obtained from separate PCAs based on data on visible ask and bid side price impacts combined with the incremental price impact brought about by hidden orders. The table reports the results of a variance decomposition of the long-run forecasting error of the mid-quote change. The information shares (in %) attributable to the four factors (sum of ask and bid side) are reported in the columns $F1$, $F2$ $F3$ and $F4$. The column Σ reports the aggregated information share of all eight factors. The information share of the mid-quote innovations itself is found in column *own*. The variance shares are averaged across all stocks within the same trading activity (measured in trades per day) quartile. The numbers in parentheses are cross-sectional standard deviations. The table shows the results for four alternative orderings for a Cholesky decomposition of the variance covariance matrix of the residuals of the VAR in standard form. The orderings interchange the position of the mid-quote change (MQ) and the first two and the last two factors. Within each ordering the ask factors are placed before the bid side factors. For example the ordering $F1-F2-F3-F4-MQ$ reads in detail $F1A-F1B-F2A-F2B-F3A-F3B-F4A-F4B-MQ$.

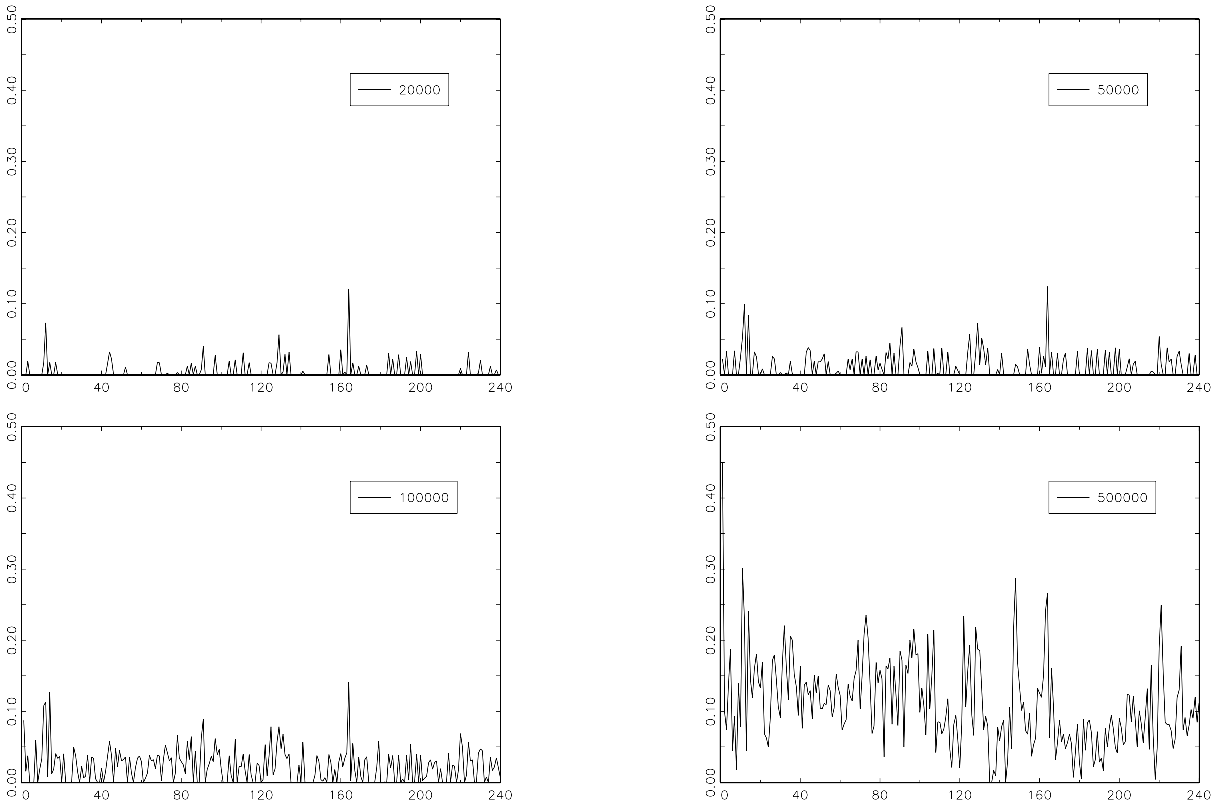


Figure 1: Time series paths for v-euros ask price impacts, complete book, BAYER. The first 240 observations at a 5-minute sampling time are plotted (102 observations per trading day).

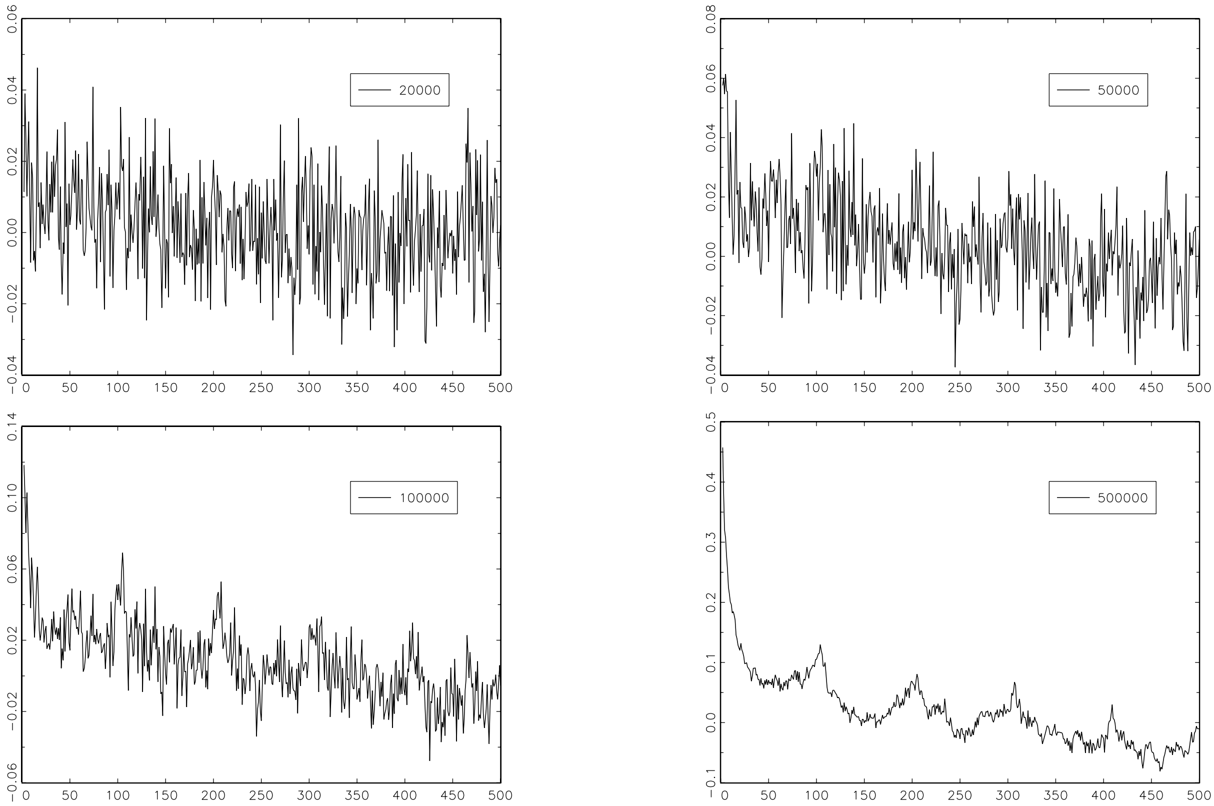


Figure 2: Autocorrelograms for v-euros ask price impacts (500 lags), complete book, BAYER. One trading day has 102 observations.

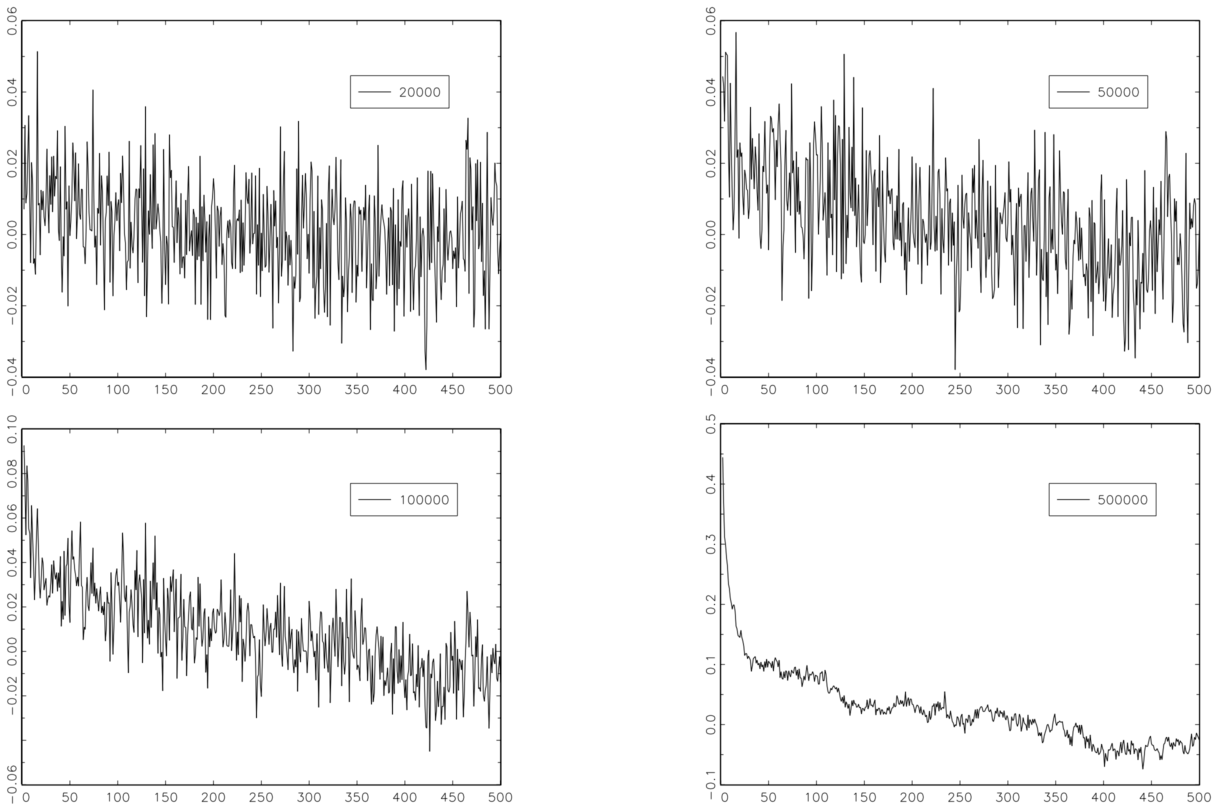


Figure 3: Autocorrelograms for v-euros ask price impacts (500 lags) normalized by their TOD mean and variance, complete book, BAYER. One trading day has 102 observations.

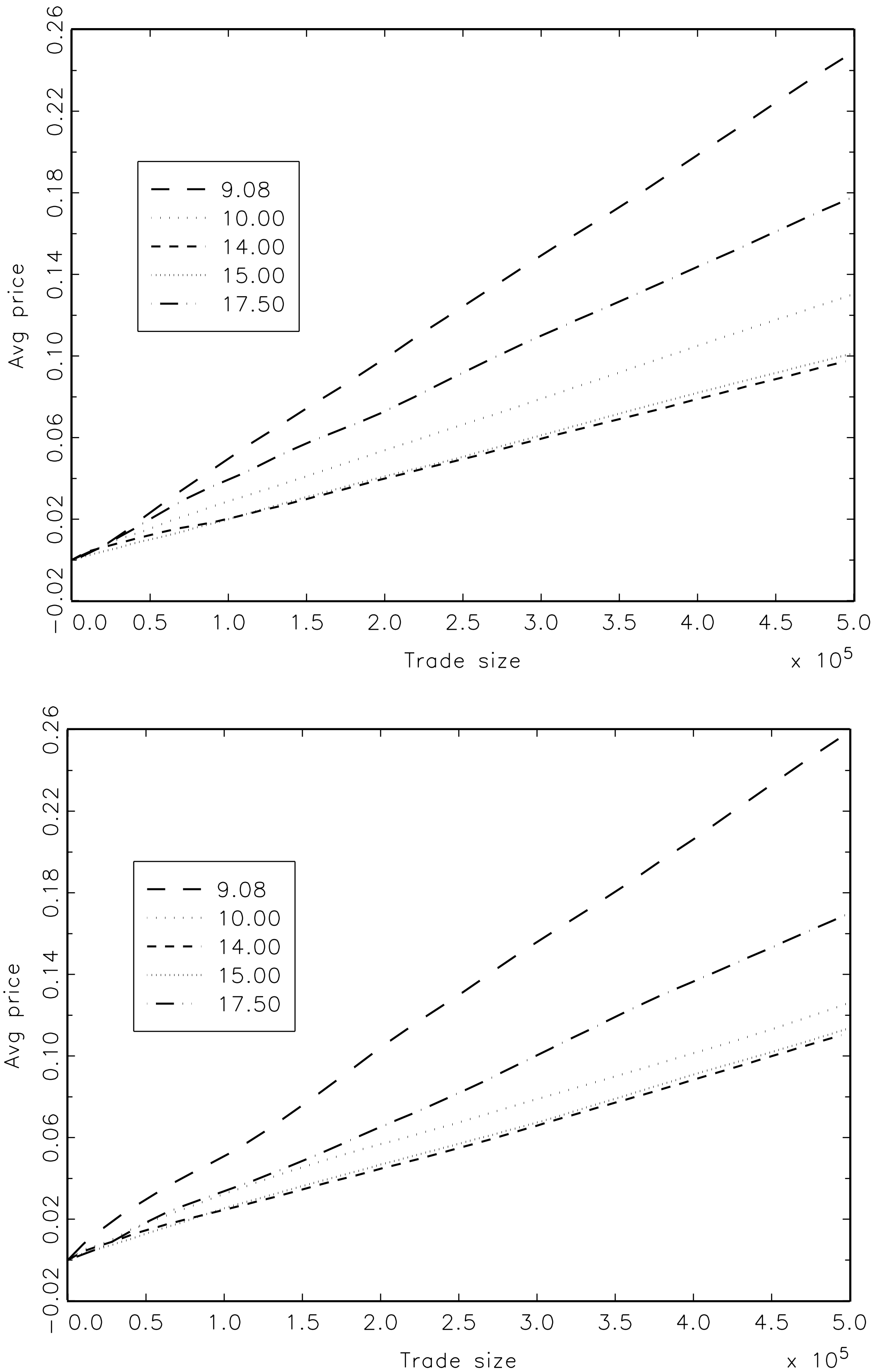
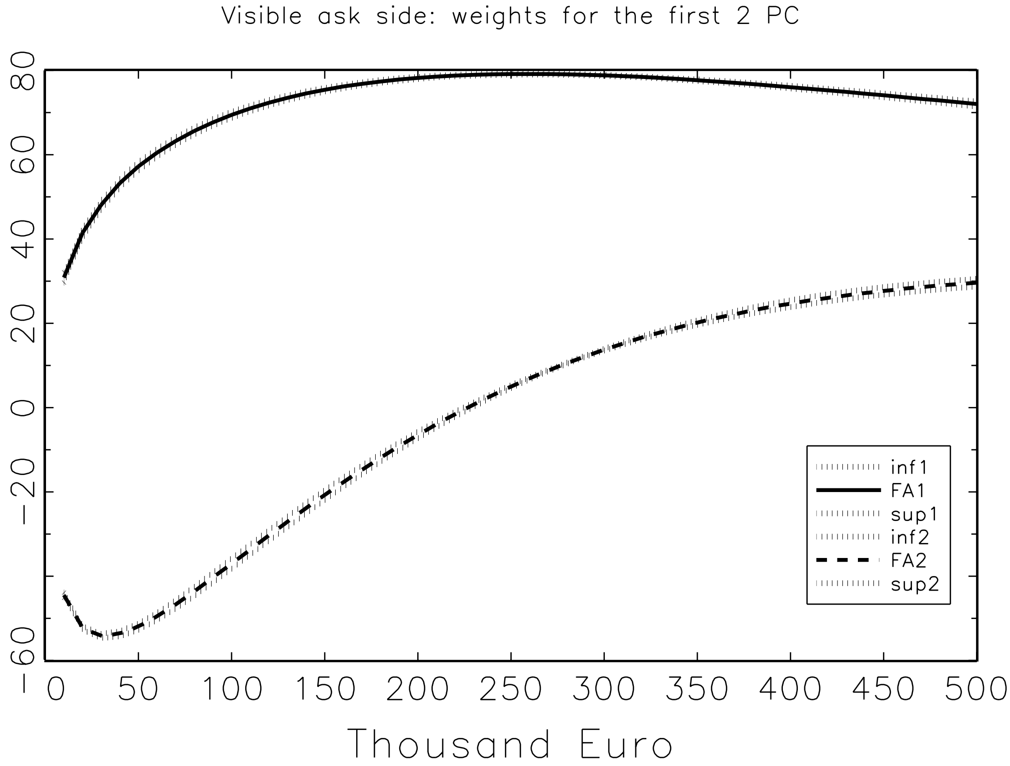


Figure 4: Time-of-day seasonality for the ask (top) and bid (bottom) price impacts.



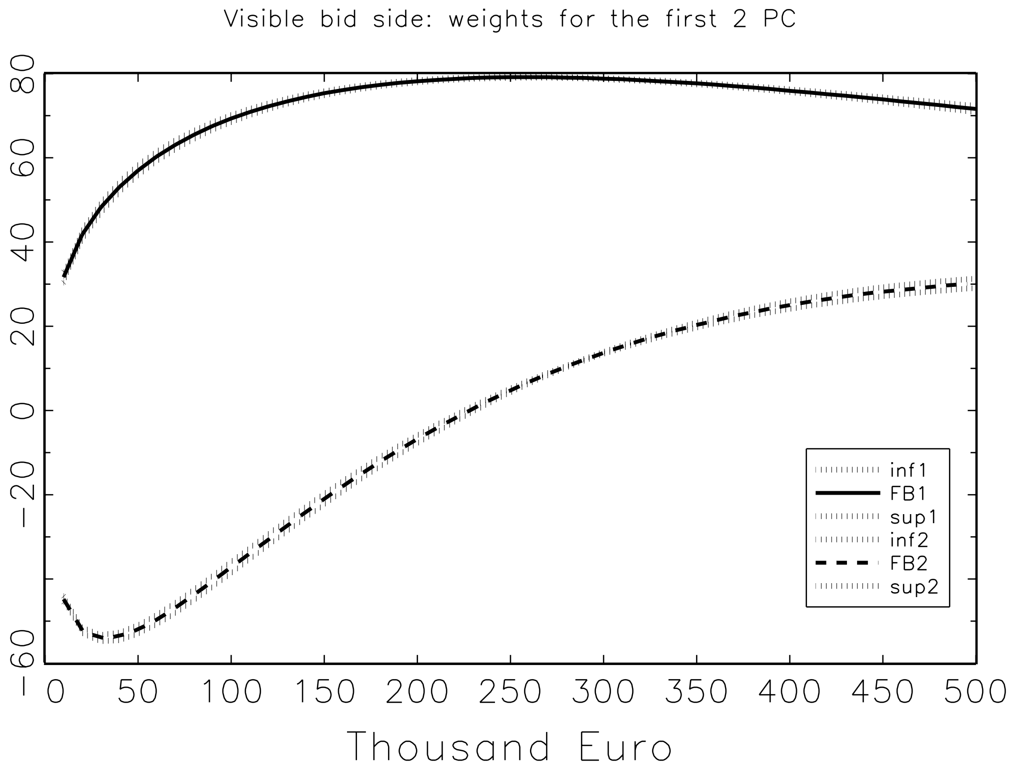


Figure 5: First and second principal components weights for the visible ask side (top of figure) and visible bid side (bottom of figure) price impacts. Averages for the 30 stocks in our sample. The dotted line are the confidence bands at 95%.

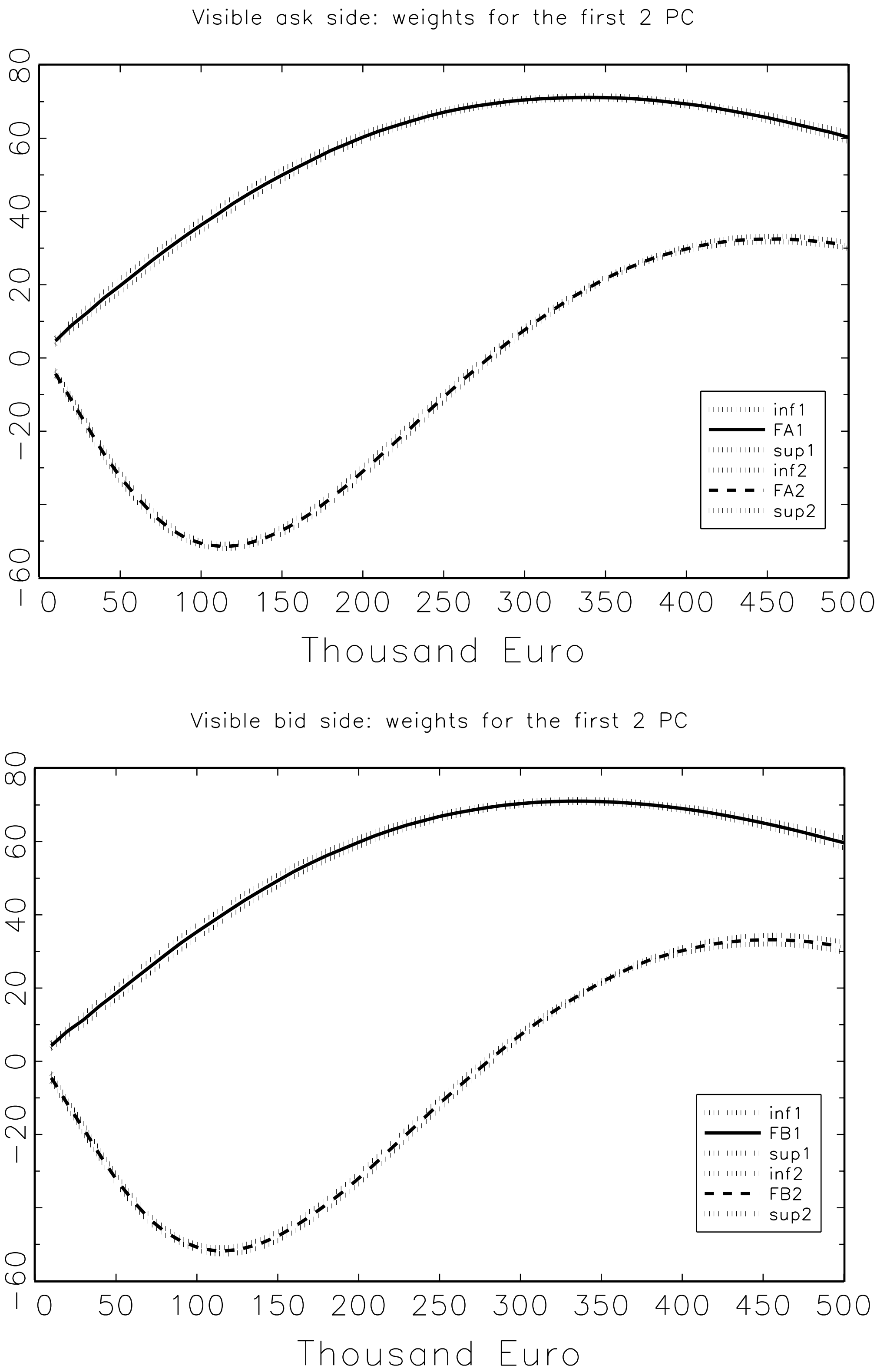
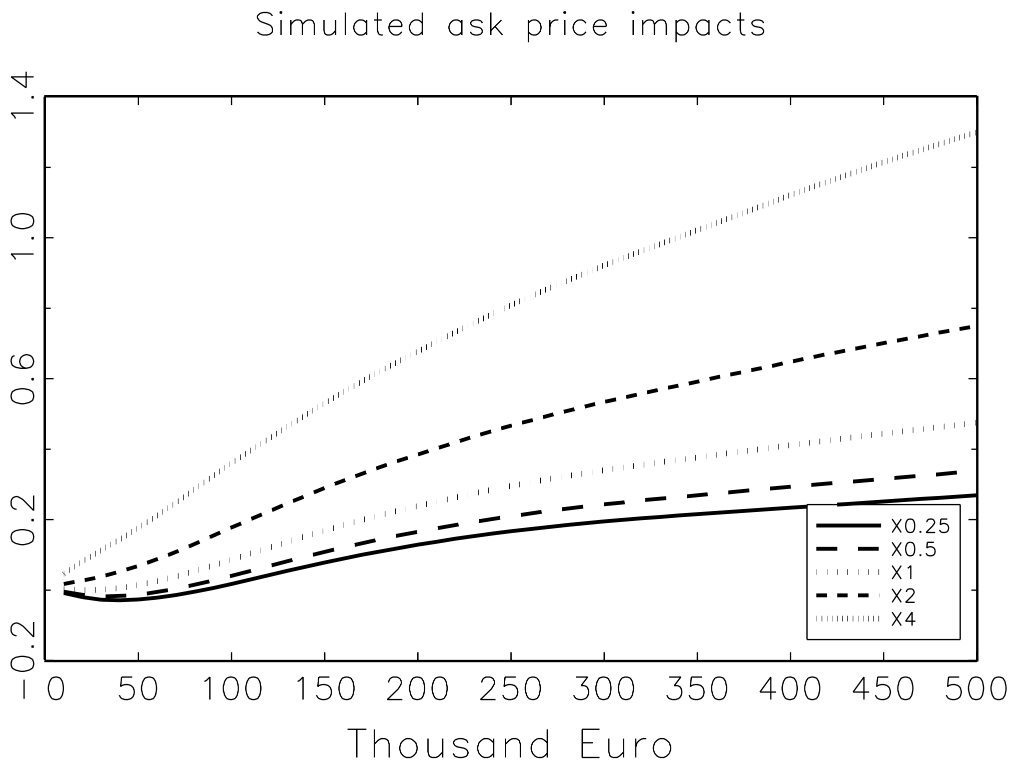


Figure 6: First and second principal components weights for the visible ask side (top of figure) and visible bid side (bottom of figure) price impact differences. Averages for the 30 stocks in our sample. The dotted line are the confidence bands at 95%.



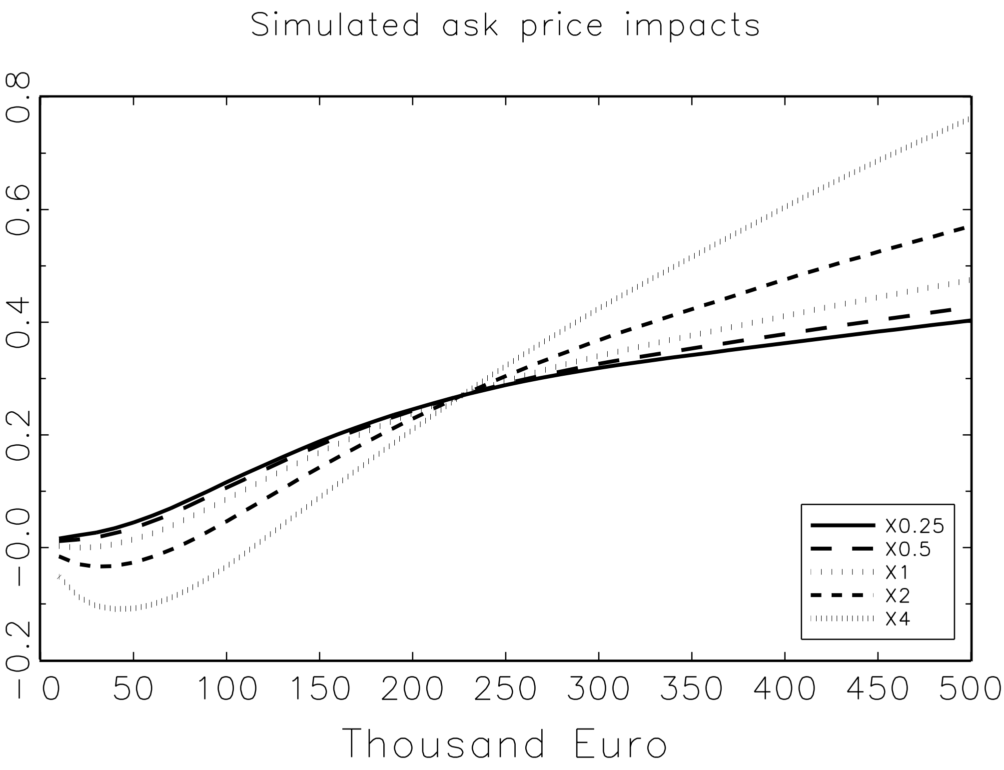


Figure 7: Simulations of the ask side of the book (Bayer stock) when the first and second principal components are increased (×2, ×4) or decreased (×0.5, ×0.25). The base scenario is characterized as ×1. The top panel shows the simulations for the first principal component, the bottom panel the simulations for the second principal components.

cfr working paper series

centre for financial research
cologne

cfr working papers are available for download from **www.cfr-cologne.de**.

Hardcopies can be ordered from: centre for financial research (cfr),
Albertus Magnus Platz, 50923 Koeln, Germany.

2012

No.	Author(s)	Title
12-06	A. Kempf, A. Pütz, F. Sonnenburg	Fund Manager Duality: Impact on Performance and Investment Behavior
12-05	R. Wermers	Runs on Money Market Mutual Funds
12-04	R. Wermers	A matter of style: The causes and consequences of style drift in institutional portfolios
12-03	C. Andres, A. Betzer, I. van den Bongard, C. Haesner, E. Theissen	Dividend Announcements Reconsidered: Dividend Changes versus Dividend Surprises
12-02	C. Andres, E. Fernau, E. Theissen	Is It Better To Say Goodbye? When Former Executives Set Executive Pay
12-01	L. Andreu, A. Pütz	Are Two Business Degrees Better Than One? Evidence from Mutual Fund Managers' Education

2011

No.	Author(s)	Title
11-16	V. Agarwal, J.-P. Gómez, R. Priestley	Management Compensation and Market Timing under Portfolio Constraints
11-15	T. Dimpfl, S. Jank	Can Internet Search Queries Help to Predict Stock Market Volatility?
11-14	P. Gomber, U. Schweickert, E. Theissen	Liquidity Dynamics in an Electronic Open Limit Order Book: An Event Study Approach
11-13	D. Hess, S. Orbe	Irrationality or Efficiency of Macroeconomic Survey Forecasts? Implications from the Anchoring Bias Test
11-12	D. Hess, P. Immenkötter	Optimal Leverage, its Benefits, and the Business Cycle
11-11	N. Heinrichs, D. Hess, C. Homburg, M. Lorenz, S. Sievers	Extended Dividend, Cash Flow and Residual Income Valuation Models – Accounting for Deviations from Ideal Conditions
11-10	A. Kempf, O. Korn, S. Saßning	Portfolio Optimization using Forward - Looking Information

11-09	V. Agarwal, S. Ray	Determinants and Implications of Fee Changes in the Hedge Fund Industry
11-08	G. Cici, L.-F. Palacios	On the Use of Options by Mutual Funds: Do They Know What They Are Doing?
11-07	V. Agarwal, G. D. Gay, L. Ling	Performance inconsistency in mutual funds: An investigation of window-dressing behavior
11-06	N. Hautsch, D. Hess, D. Veredas	The Impact of Macroeconomic News on Quote Adjustments, Noise, and Informational Volatility
11-05	G. Cici	The Prevalence of the Disposition Effect in Mutual Funds' Trades
11-04	S. Jank	Mutual Fund Flows, Expected Returns and the Real Economy
11-03	G.Fellner, E.Theissen	Short Sale Constraints, Divergence of Opinion and Asset Value: Evidence from the Laboratory
11-02	S.Jank	Are There Disadvantaged Clienteles in Mutual Funds?
11-01	V. Agarwal, C. Meneghetti	The Role of Hedge Funds as Primary Lenders

2010

No.	Author(s)	Title
10-20	G. Cici, S. Gibson, J.J. Merrick Jr.	Missing the Marks? Dispersion in Corporate Bond Valuations Across Mutual Funds
10-19	J. Hengelbrock, E. Theissen, C. Westheide	Market Response to Investor Sentiment
10-18	G. Cici, S. Gibson	The Performance of Corporate-Bond Mutual Funds: Evidence Based on Security-Level Holdings
10-17	D. Hess, D. Kreutzmann, O. Pucker	Projected Earnings Accuracy and the Profitability of Stock Recommendations
10-16	S. Jank, M. Wedow	Sturm und Drang in Money Market Funds: When Money Market Funds Cease to Be Narrow
10-15	G. Cici, A. Kempf, A. Puetz	The Valuation of Hedge Funds' Equity Positions
10-14	J. Grammig, S. Jank	Creative Destruction and Asset Prices
10-13	S. Jank, M. Wedow	Purchase and Redemption Decisions of Mutual Fund Investors and the Role of Fund Families
10-12	S. Artmann, P. Finter, A. Kempf, S. Koch, E. Theissen	The Cross-Section of German Stock Returns: New Data and New Evidence
10-11	M. Chesney, A. Kempf	The Value of Tradeability
10-10	S. Frey, P. Herbst	The Influence of Buy-side Analysts on Mutual Fund Trading
10-09	V. Agarwal, W. Jiang, Y. Tang, B. Yang	Uncovering Hedge Fund Skill from the Portfolio Holdings They Hide
10-08	V. Agarwal, V. Fos, W. Jiang	Inferring Reporting Biases in Hedge Fund Databases from Hedge Fund Equity Holdings
10-07	V. Agarwal, G. Bakshi,	Do Higher-Moment Equity Risks Explain Hedge Fund

	J. Huij	Returns?
10-06	J. Grammig, F. J. Peter	Tell-Tale Tails
10-05	K. Drachter, A. Kempf	Höhe, Struktur und Determinanten der Managervergütung- Eine Analyse der Fondsbranche in Deutschland
10-04	J. Fang, A. Kempf, M. Trapp	Fund Manager Allocation
10-03	P. Finter, A. Niessen-Ruenzi, S. Ruenzi	The Impact of Investor Sentiment on the German Stock Market
10-02	D. Hunter, E. Kandel, S. Kandel, R. Wermers	Endogenous Benchmarks
10-01	S. Artmann, P. Finter, A. Kempf	Determinants of Expected Stock Returns: Large Sample Evidence from the German Market

2009

No.	Author(s)	Title
09-17	E. Theissen	Price Discovery in Spot and Futures Markets: A Reconsideration
09-16	M. Trapp	Trading the Bond-CDS Basis – The Role of Credit Risk and Liquidity
09-15	A. Betzer, J. Gider, D.Metzger, E. Theissen	Strategic Trading and Trade Reporting by Corporate Insiders
09-14	A. Kempf, O. Korn, M. Uhrig-Homburg	The Term Structure of Illiquidity Premia
09-13	W. Bühler, M. Trapp	Time-Varying Credit Risk and Liquidity Premia in Bond and CDS Markets
09-12	W. Bühler, M. Trapp	Explaining the Bond-CDS Basis – The Role of Credit Risk and Liquidity
09-11	S. J. Taylor, P. K. Yadav, Y. Zhang	Cross-sectional analysis of risk-neutral skewness
09-10	A. Kempf, C. Merkle, A. Niessen-Ruenzi	Low Risk and High Return – Affective Attitudes and Stock Market Expectations
09-09	V. Fotak, V. Raman, P. K. Yadav	Naked Short Selling: The Emperor`s New Clothes?
09-08	F. Bardong, S.M. Bartram, P.K. Yadav	Informed Trading, Information Asymmetry and Pricing of Information Risk: Empirical Evidence from the NYSE
09-07	S. J. Taylor , P. K. Yadav, Y. Zhang	The information content of implied volatilities and model-free volatility expectations: Evidence from options written on individual stocks
09-06	S. Frey, P. Sandas	The Impact of Iceberg Orders in Limit Order Books
09-05	H. Beltran-Lopez, P. Giot, J. Grammig	Commonalities in the Order Book
09-04	J. Fang, S. Ruenzi	Rapid Trading bei deutschen Aktienfonds: Evidenz aus einer großen deutschen Fondsgesellschaft
09-03	A. Banegas, B. Gillen, A. Timmermann, R. Wermers	The Performance of European Equity Mutual Funds

09-02	J. Grammig, A. Schrimpf, M. Schuppli	Long-Horizon Consumption Risk and the Cross-Section of Returns: New Tests and International Evidence
09-01	O. Korn, P. Koziol	The Term Structure of Currency Hedge Ratios

2008

No.	Author(s)	Title
08-12	U. Bonenkamp, C. Homburg, A. Kempf	Fundamental Information in Technical Trading Strategies
08-11	O. Korn	Risk Management with Default-risky Forwards
08-10	J. Grammig, F.J. Peter	International Price Discovery in the Presence of Market Microstructure Effects
08-09	C. M. Kuhnen, A. Niessen	Public Opinion and Executive Compensation
08-08	A. Pütz, S. Ruenzi	Overconfidence among Professional Investors: Evidence from Mutual Fund Managers
08-07	P. Osthoff	What matters to SRI investors?
08-06	A. Betzer, E. Theissen	Sooner Or Later: Delays in Trade Reporting by Corporate Insiders
08-05	P. Linge, E. Theissen	Determinanten der Aktionärspräsenz auf Hauptversammlungen deutscher Aktiengesellschaften
08-04	N. Hautsch, D. Hess, C. Müller	Price Adjustment to News with Uncertain Precision
08-03	D. Hess, H. Huang, A. Niessen	How Do Commodity Futures Respond to Macroeconomic News?
08-02	R. Chakrabarti, W. Megginson, P. Yadav	Corporate Governance in India
08-01	C. Andres, E. Theissen	Setting a Fox to Keep the Geese - Does the Comply-or-Explain Principle Work?

2007

No.	Author(s)	Title
07-16	M. Bär, A. Niessen, S. Ruenzi	The Impact of Work Group Diversity on Performance: Large Sample Evidence from the Mutual Fund Industry
07-15	A. Niessen, S. Ruenzi	Political Connectedness and Firm Performance: Evidence From Germany
07-14	O. Korn	Hedging Price Risk when Payment Dates are Uncertain
07-13	A. Kempf, P. Osthoff	SRI Funds: Nomen est Omen
07-12	J. Grammig, E. Theissen, O. Wuensche	Time and Price Impact of a Trade: A Structural Approach
07-11	V. Agarwal, J. R. Kale	On the Relative Performance of Multi-Strategy and Funds of Hedge Funds
07-10	M. Kasch-Haroutounian, E. Theissen	Competition Between Exchanges: Euronext versus Xetra
07-09	V. Agarwal, N. D. Daniel, N. Y. Naik	Do hedge funds manage their reported returns?

07-08	N. C. Brown, K. D. Wei, R. Wermers	Analyst Recommendations, Mutual Fund Herding, and Overreaction in Stock Prices
07-07	A. Betzer, E. Theissen	Insider Trading and Corporate Governance: The Case of Germany
07-06	V. Agarwal, L. Wang	Transaction Costs and Value Premium
07-05	J. Grammig, A. Schrimpf	Asset Pricing with a Reference Level of Consumption: New Evidence from the Cross-Section of Stock Returns
07-04	V. Agarwal, N.M. Boyson, N.Y. Naik	Hedge Funds for retail investors? An examination of hedged mutual funds
07-03	D. Hess, A. Niessen	The Early News Catches the Attention: On the Relative Price Impact of Similar Economic Indicators
07-02	A. Kempf, S. Ruenzi, T. Thiele	Employment Risk, Compensation Incentives and Managerial Risk Taking - Evidence from the Mutual Fund Industry -
07-01	M. Hagemeister, A. Kempf	CAPM und erwartete Renditen: Eine Untersuchung auf Basis der Erwartung von Marktteilnehmern

2006

No.	Author(s)	Title
06-13	S. Čeljo-Hörhager, A. Niessen	How do Self-fulfilling Prophecies affect Financial Ratings? - An experimental study
06-12	R. Wermers, Y. Wu, J. Zechner	Portfolio Performance, Discount Dynamics, and the Turnover of Closed-End Fund Managers
06-11	U. v. Lilienfeld-Toal, S. Ruenzi	Why Managers Hold Shares of Their Firm: An Empirical Analysis
06-10	A. Kempf, P. Osthoff	The Effect of Socially Responsible Investing on Portfolio Performance
06-09	R. Wermers, T. Yao, J. Zhao	Extracting Stock Selection Information from Mutual Fund holdings: An Efficient Aggregation Approach
06-08	M. Hoffmann, B. Kempa	The Poole Analysis in the New Open Economy Macroeconomic Framework
06-07	K. Drachter, A. Kempf, M. Wagner	Decision Processes in German Mutual Fund Companies: Evidence from a Telephone Survey
06-06	J.P. Krahnen, F.A. Schmid, E. Theissen	Investment Performance and Market Share: A Study of the German Mutual Fund Industry
06-05	S. Ber, S. Ruenzi	On the Usability of Synthetic Measures of Mutual Fund Net-Flows
06-04	A. Kempf, D. Mayston	Liquidity Commonality Beyond Best Prices
06-03	O. Korn, C. Koziol	Bond Portfolio Optimization: A Risk-Return Approach
06-02	O. Scaillet, L. Barras, R. Wermers	False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas
06-01	A. Niessen, S. Ruenzi	Sex Matters: Gender Differences in a Professional Setting

2005

No.	Author(s)	Title
05-16	E. Theissen	An Analysis of Private Investors´ Stock Market Return Forecasts

05-15	T. Foucault, S. Moinas, E. Theissen	Does Anonymity Matter in Electronic Limit Order Markets
05-14	R. Kosowski, A. Timmermann, R. Wermers, H. White	Can Mutual Fund „Stars“ Really Pick Stocks? New Evidence from a Bootstrap Analysis
05-13	D. Avramov, R. Wermers	Investing in Mutual Funds when Returns are Predictable
05-12	K. Griese, A. Kempf	Liquiditätsdynamik am deutschen Aktienmarkt
05-11	S. Ber, A. Kempf, S. Ruenzi	Determinanten der Mittelzuflüsse bei deutschen Aktienfonds
05-10	M. Bär, A. Kempf, S. Ruenzi	Is a Team Different From the Sum of Its Parts? Evidence from Mutual Fund Managers
05-09	M. Hoffmann	Saving, Investment and the Net Foreign Asset Position
05-08	S. Ruenzi	Mutual Fund Growth in Standard and Specialist Market Segments
05-07	A. Kempf, S. Ruenzi	Status Quo Bias and the Number of Alternatives - An Empirical Illustration from the Mutual Fund Industry
05-06	J. Grammig, E. Theissen	Is Best Really Better? Internalization of Orders in an Open Limit Order Book
05-05	H. Beltran-Lopez, J. Grammig, A.J. Menkveld	Limit order books and trade informativeness
05-04	M. Hoffmann	Compensating Wages under different Exchange rate Regimes
05-03	M. Hoffmann	Fixed versus Flexible Exchange Rates: Evidence from Developing Countries
05-02	A. Kempf, C. Memmel	Estimating the Global Minimum Variance Portfolio
05-01	S. Frey, J. Grammig	Liquidity supply and adverse selection in a pure limit order book market

2004

No.	Author(s)	Title
04-10	N. Hautsch, D. Hess	Bayesian Learning in Financial Markets – Testing for the Relevance of Information Precision in Price Discovery
04-09	A. Kempf, K. Kreuzberg	Portfolio Disclosure, Portfolio Selection and Mutual Fund Performance Evaluation
04-08	N.F. Carline, S.C. Linn, P.K. Yadav	Operating performance changes associated with corporate mergers and the role of corporate governance
04-07	J.J. Merrick, Jr., N.Y. Naik, P.K. Yadav	Strategic Trading Behaviour and Price Distortion in a Manipulated Market: Anatomy of a Squeeze
04-06	N.Y. Naik, P.K. Yadav	Trading Costs of Public Investors with Obligatory and Voluntary Market-Making: Evidence from Market Reforms
04-05	A. Kempf, S. Ruenzi	Family Matters: Rankings Within Fund Families and Fund Inflows
04-04	V. Agarwal, N.D. Daniel, N.Y. Naik	Role of Managerial Incentives and Discretion in Hedge Fund Performance
04-03	V. Agarwal, W.H. Fung, J.C. Loon, N.Y. Naik	Risk and Return in Convertible Arbitrage: Evidence from the Convertible Bond Market
04-02	A. Kempf, S. Ruenzi	Tournaments in Mutual Fund Families

04-01	I. Chowdhury, M. Hoffmann, A. Schabert	Inflation Dynamics and the Cost Channel of Monetary Transmission

centre for financial research
cfr/university of cologne
Albertus-Magnus-Platz
D-50923 cologne
fon +49(0)221-470-6995
fax +49(0)221-470-3992
kempf@cfr-cologne.de
www.cfr-cologne.de