CENTER FOR FINANCIAL STUDIES

No. 2011/09

Limit Order Books and Trade Informativeness

Hélena Beltran-Lopez, Joachim Grammig,
and Albert J. Menkveld

Center for Financial Studies

The *Center for Financial Studies* is a nonprofit research organization, supported by an association of more than 120 banks, insurance companies, industrial corporations and public institutions. Established in 1968 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia.

The CFS Working Paper Series presents the result of scientific research on selected topics in the field of money, banking and finance. The authors were either participants in the Center´s Research Fellow Program or members of one of the Center´s Research Projects.

If you would like to know more about the *Center for Financial Studies*, please let us know of your interest.

Prof. Michalis Haliassos, Ph.D.   Prof. Dr. Jan Pieter Krahnen   Prof. Dr. Uwe Walz

CFS Working Paper No. 2011/09

Limit Order Books and Trade Informativeness

Hélena Beltran-Lopez[1], Joachim Grammig[2],
and Albert J. Menkveld[3],

January 30, 2011

Abstract:
In the microstructure literature, information asymmetry is an important determinant of market liquidity. The classic setting is that uninformed dedicated liquidity suppliers charge price concessions when incoming market orders are likely to be informationally motivated. In limit order book markets, however, this relationship is less clear, as market participants can switch roles, and freely choose to immediately demand or patiently supply liquidity by submitting either market or limit orders. We study the importance of information asymmetry in limit order books based on a recent sample of thirty German DAX stocks. We find that Hasbrouck's (1991) measure of trade informativeness Granger-causes book liquidity, in particular that required to fill large market orders. Picking-off risk due to public news induced volatility is more important for top-of-the book liquidity supply. In our multivariate analysis we control for volatility, trading volume, trading intensity and order imbalance to isolate the effect of trade informativeness on book liquidity.

1 Fifth Third Asset Management, Inc., 600 Superior Ave. East, MD A6516A, Cleveland OH 44114, U.S.A.;
  E-mail: helena.beltran-lopez@53.com

2 Eberhard-Karls-University Tübingen and Center for Financial Studies. Contact Address: Eberhard-Karls-University Tübingen, Mohlstr. 36, 72074 Tübingen, Germany; E-mail: joachim.grammig@uni-tuebingen.de

3 VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands; E-mail: ajmenkveld@feweb.vu.nl

1   Introduction

The classic microstructure literature distinguishes liquidity suppliers and liquidity demanders, which naturally introduces information asymmetry. That is, liquidity suppliers trade against potentially privately-informed liquidity demanders and charge them an increased price concession to protect themselves.[1] This deters uninformed, hedging-motivated liquidity demand and, in the extreme, might cause the market to break down. Information asymmetry thus reduces welfare (cf. Biais et al. 2005, p.223-227). Easley et al. (2002) provide evidence that asymmetric information risk is priced, as stocks for which they estimate a high probability of informed trading have to offer higher expected returns.

With the advent of electronic limit order book (LOB) markets, however, the distinction between uninformed liquidity suppliers and potentially informed liquidity demanders became blurred. Investors arriving at the market can choose to demand liquidity through a market order, but they can also enter their trading interest in the book via a limit order. In the latter case they effectively supply liquidity. It is therefore unclear to what extent the increased price concession due to information asymmetry, one of the cornerstones of classic microstructure, still matters for liquidity supply in electronic LOB markets.

We exploit a comprehensive sample of thirty index stocks traded in the limit order book of the German Stock Exchange to empirically assess the effect of information asymmetry on the supply of liquidity. The main advantage is that the data come from a pure limit order book market which, for these stocks, captures over 95% of the non-OTC order flow.

In a time series approach we relate Hasbrouck's (1991) informativeness measure to limit order book liquidity. We find that trade informativeness Granger-causes price concessions for large market orders, but has little impact for average-size market orders. For top-of-the book liquidity supply, picking-off risk due to public news induced volatility is more important. In a multivariate analysis of the impact of trade informativeness on book liquidity we allow for control variables such as realised volatility, trading intensity, and trade size.

The main motivation for our study is to provide empirical evidence to feed the rapidly expanding theoretical literature on limit order markets. Recent LOB theory can broadly be classified into static and dynamic models. Static models strictly distinguish between liquidity suppliers and liquidity demanders. Only the latter have access to private information about the fundamental asset price. The limit order book is the optimal market structure in this framework, as it fosters competition among suppliers of liquidity. In par-

ticular, Glosten (1994) shows that risk-neutral limit order submitters compete for supply so that, in equilibrium, the marginal order breaks even in terms of expected profit. Limit order traders incur adverse-selection costs as they trade against a (potentially) informed market order. Biais et al. (2000) consider an extension with a limited number of strategic suppliers. They show that the Glosten result obtains if their number goes to infinity. Seppi (1997) and Parlour and Seppi (2003) compare the LOB market with a hybrid market, where the limit order book competes with a strategic specialist who has the privilege of ex-post price improvement. Biais et al. (1998) show how the LOB market is more likely to implement the competitive equilibrium of strategic suppliers when compared to a dealer or a floor market.

The class of dynamic LOB models does not distinguish ex-ante between liquidity suppliers and demanders, but let agents arrive randomly in the market to decide whether to submit a limit order (act as liquidity supplier) or to submit a market order (act as liquidity demander), or do nothing. These agents trade-off the cost of immediacy (to pay the spread) associated with a market order against the costs of a limit order submission, i.e. possibly infinitely delayed execution and picking-off risk. Picking-off risk occurs when limit orders are not monitored continuously, so that public information arrivals mechanically make limit buys execute more often when the value drops and make limit sells execute more often when the value rises. The agents in these dynamic models trade to lock in some private value orthogonal to common value innovations. Information is symmetric and there is no adverse-selection risk (cf. Parlour 1998, Foucault 1999, Goettler et al. 2005, Hollifield et al. 2006, Hollifield et al. 2004, Foucault et al. 2005, Rosu 2009). To the best of our knowledge, Goettler et al. (2009) is the single exception as they propose a dynamic LOB model with adverse-selection risk.

One reason why dynamic LOB models abstract from adverse-selection risk is mathematical tractability, but Foucault et al. (2005) and Rosu (2009) also justify the assumption with Huang and Stoll's (1997) finding that the majority of the bid-ask spread (88.8%) is due to non-informational frictions. We note that while dynamic LOB models do consider picking-off risk, this risk is fundamentally different from adverse-selection risk, as the latter involves a (potential) transfer of surplus from uninformed to informed traders and, therefore, potentially impedes trade. We henceforth interpret picking-off risk in the narrow sense of "adverse" execution due to public information, consistent with dynamic LOB models. In real-world markets, such adverse execution might also be due to private information in the order flow. In our multivariate analysis, both aspects of adverse execution are accounted for by including a proxy for public news induced volatility (realised volatility) as a control variable along with Hasbrouck's (1991) measure of trade informativeness.

In addition to the time series analysis of trade informativeness and liquidity supply, we also pursue a structural approach in order to test the Glosten (1994) model. For this test, we have a limit order book in mind that quickly replenishes to equilibrium after each market order. We follow Sandås (2001) who tests the Glosten model using three months of data (starting Dec 1991) for ten stocks traded on the Stockholm Stock Exchange (SSE). Formal tests performed by Sandås (2001) reject the model, but one could argue that the SSE market design is too different from the theoretical setting of the Glosten model in the first place. We believe our data are more appropriate because of three main reasons. First, the German electronic market covers over 95% of non-OTC trades, whereas the Swedish sample missed "a significant fraction of the turnover" due to transactions on London's SEAQ International and NASDAQ (see Sandås (2001, p.708)). Second, entry as an implicit market maker through limit order activity is more attractive. The reason is that the German exchange, as opposed to the 1991 Stockholm exchange, does not charge for limit order submission or cancellation. Third, our more recent sample benefits from technological development since 1991. Early evidence of a quickly replenishing book is the high limit order activity that we measure in our data. We find that the ratio of limit to market orders is 6.1, which compares to a 1.7 ratio for Sandås' SSE sample.

Empirical evidence consistent with a replenishing "Glosten"-book is that on poor book liquidity limit orders are more likely than market orders (cf. Biais et al. 1995, Griffiths et al. 2000, Ahn et al. 2001, Ranaldo 2004). This evidence, however, is also consistent with dynamic limit order book models. There is, however, also some evidence in favour of a replenishing Glosten-book after *privately informed* market orders, which is harder to reconcile with current dynamic models that assume symmetric information. As for private information in market orders, Biais et al. (1995) study the Paris LOB market and find that, along with the ask, the bid changes after a market buy, which indicates that market orders are informative. Bloomfield et al. (2009) conduct an LOB market experiment and one of their findings is that informed traders use market orders relatively more often than noise traders or liquidity traders. As for a quickly replenishing book, Biais et al. (1995, p.1693) document that durations are 30% shorter after a large market order and interpret this finding as "traders quickly place orders within the best quotes to supply liquidity at relatively advantageous prices and to obtain time priority (p.1683)". Ranaldo (2004) reports a similar finding for the Swiss Stock Exchange. Finally, LOB markets are easily accessible, transparent electronic markets, and any profit opportunities in the book should therefore be quickly filled by outside liquidity providers.

There is also evidence against a quickly replenishing Glosten-LOB. For the hybrid

NYSE market, Harris and Hasbrouck (1996, p.230) find that "expected profits accruing to an off-the-floor trader who attempts to behave as a dealer are generally negative...". In hybrid markets, though, limit orders are at a disadvantage relative to the specialist who can cream-skim arriving market orders, as she can decide ex-post (i.e. after arrival) whether or not to supply liquidity (cf. Seppi 1997, Parlour and Seppi 2003).

Our main results can be summarised as follows. Consistent with Sandås's (2001) results, our formal tests also reject the structural model. The informativeness parameter, Glosten's α, turns insignificant when based on the updating conditions for an equilibrium LOB. One interpretation is that book replenishment from transaction to transaction is noisy and potentially incomplete. This motivates our alternative approach that does not rely on a parametrised model, and uses the long-term price impact of trades to measure their informativeness. This approach finds empirical support for the main prediction of static LOB models that informativeness matters for book liquidity, which we measure through price concessions of market orders of different sizes. We find that large order price concessions are most sensitive to trade informativeness, as opposed to any of the control variables employed in the multivariate analysis, e.g. realised volatility. However, we also find that the bid-ask spread and the average-size order price concessions respond stronger to our proxy of market volatility, i.e. picking-off risk that is unrelated to private information.

These findings relate to two recent papers on the Island ECN, which has the additional feature of competition for order flow with other non-OTC venues, most notably the NASDAQ. Consistent with our findings, Hasbrouck and Saar (2007) document for a cross-section of securities that volatility is associated with lower depth in the book. We contribute by emphasising trade informativeness, where we control for volatility. And, Hasbrouck and Saar (2009) find that limit orders should not be viewed solely as "patient" liquidity supply, as some "fleeting" limit orders appear to hunt for hidden depth.

The remainder of the paper is organised as follows. Section 2 discusses the institutional background, the available data, and presents summary statistics. Section 3 reviews the Glosten (1994) model and estimates the structural parameters by GMM following Sandås (2001). Section 4 studies time-varying trade informativeness using the Hasbrouck (1991) measure and conducts Granger causality tests to study whether high informativeness causes poor book liquidity and vice versa. In a multivariate analysis we add control variables such as volatility, trade size, and duration to isolate the trade informativeness effect. Section 5 discusses the results and Section 6 concludes.

2 Institutional Background and Data Sample

2.1 The XETRA Limit Order Book

The German Stock Exchange (GSE) operates the electronic limit order book XETRA according to trading rules that are similar to previously studied limit order markets e.g. Euronext, the Hong Kong Stock Exchange and the Swedish Stock Exchange. We refer to Biais et al. (1995) for a detailed description of these rules. Trading starts at 9 a.m. (Central European Time CET) with an opening auction and closes at 5.30 p.m. with a closing auction. Around noon, trading is interrupted for the "mid-day" auction. For the DAX30 constituent stocks studied in this paper there are no designated market makers. Market orders larger than the depth available at the best quote automatically "walk up the book."

XETRA is quite close to the stylised setting analysed in Glosten's (1994) limit order book model, but deviates in two ways. First, it allows for so-called iceberg orders, which are similar to standard limit orders with the exception that part of the limit order volume is not displayed in the book. This hidden volume enjoys price priority over other limit orders, but not time priority. In the remainder of the paper, we focus on the results based on the visible book. For the sake of a robustness check, we also perform the analysis on the total book and find that the results are not affected. The second main difference to the Glosten framework is that the XETRA limit order book faces some local, regional, and international competition for order flow. Parallel to the XETRA system, the German Stock Exchange maintains a trading floor which, by all means, functions as an upstairs market. Regional competition comes from smaller German exchanges. Finally, eleven of the thirty DAX stocks are cross-listed as ADR at the NYSE. However, we can safely ignore these alternative trading venues in the analysis as the XETRA system has a market share of at least 95%. Our stocks also trade in an OTC market, but it is hard to measure the size of this market as trades do not need to be reported.

2.2 Data and Summary Statistics

From the GSE, we have received data on all XETRA order book events—entries, cancellations, revisions, expirations, partial-fills and full-fills of market, limit, and iceberg orders—for a three month period: January 2nd through March 31st, 2004. In this study, we focus on the thirty blue chip stocks in the German DAX index. We use the data to perform a real-time

reconstruction of the order book sequences. For that purpose, we start with an initial state of the book each day and use all order events to re-build the book sequences for the remainder of the day, accounting for every event that changes the order book. Our reconstruction procedure permits distinguishing the visible and the hidden part of the order book.

[insert Table 1 here]

Summary statistics in Table 1 show that XETRA trading is very active, in particular in terms of limit order submissions. For the average stock in the sample, the average daily number of limit orders submitted to the system is 12,785, i.e. 25 limit orders per minute. Of these submissions, 10,887 are cancelled prior to execution on the same day. The average number of trades per day is 2,099, i.e. 4 trades per minute. These numbers suggest that limit order traders actively follow the market and submit orders to benefit from profit opportunities in the spirit of the Glosten-model. It is interesting to note that relatively many orders arrive as limit orders (as opposed to market orders or marketable limit orders) in comparison to what we know from previous studies. For the XETRA data, the limit to market order ratio is 6.1, whereas this ratio is, for example, 1.1 for the Paris Bourse in November 1991 and 1.7 for the Stockholm Stock Exchange (SSE) from December 1991 through February 1992 (cf. Biais et al. 1995, Sandås 2001).

The descriptive statistics reveal three more interesting stylised facts of the data. First, we present the frequency of market orders that execute not only at the best quote, but also at prices strictly inside the book. We find that 15.2% of all market orders "walk up the book," which demonstrates the relevancy of liquidity supply beyond the best quotes. Second, we find that bid-ask spreads are small, 9 basis points on average, which is consistent with a very active and liquid market. Third, we find considerable cross-sectional variation and thus decide to sort the sample stocks into quartiles based on trade activity. Earlier work shows that informed trading is more important for small, less active stocks (cf. Easley et al. 1996).

3   Structural Econometrics: The Glosten Model

To study the importance of informativeness for book liquidity we first follow Sandås (2001) and estimate the structural parameters of the Glosten (1994) model. We briefly review the model and the implied moment conditions used for GMM estimation, and then present our

estimates.

3.1 General Features of the Model

There are two types of agents in the market: liquidity suppliers and liquidity demanders. Liquidity suppliers are patient uninformed risk-neutral agents who submit limit orders to the book to maximise expected profit. Liquidity demanders submit market orders and may have private information about the value of the security. Liquidity demanders arrive randomly at the market and we organise our time line accordingly. The sequencing of events is such that between the arrival of market orders, the book is quickly updated by liquidity suppliers, as illustrated in the following graph:
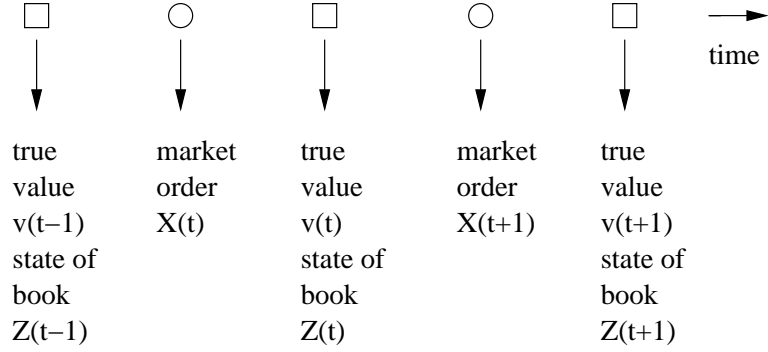
where X_t is the signed market order size (number of shares), v_t is the true value of the security after arrival of market order X_t, and Z_t captures the state of the order book (e.g., bid-ask spread, depth at the best quote).

Liquidity demanders. The liquidity demander who arrives at time t submits a market order of size X_t, a number that is positive for buys and negative for sells. We assume that buys and sells are equally likely and X_t is independent from X_s for $s \neq t$. For order size, we assume a symmetric, two-sided exponential distribution:

$$f(|X_t|) = \frac{1}{2\lambda} e^{\frac{|X_t|}{\lambda}} \tag{1}$$

where $\lambda > 0$ is the average order size in absolute terms. To capture trade informativeness, we assume the following process for the true value of the security v_t right after the arrival

of the market order at event time t:

$$v_t = E[v_t|v_{t-1}, X_t] + \eta_t = c + v_{t-1} + \alpha X_t + \eta_t \tag{2}$$

where η_t accounts for the arrival of public information between the trades at event time t and at $t-1$. The key parameter in the model is α, which captures the informativeness (with respect to the asset value) of the arriving market order. We refer to the parameter as Glosten's alpha.

Liquidity suppliers. We assume that liquidity suppliers incur fixed order-processing cost γ when transacting with incoming market orders. They submit limit orders right after the arrival of market order X_t at various prices until the marginal order at each of these prices breaks even. For example, suppose that price level $p_{1,t}$ is the lowest price above v_t at which it is profitable to supply a limit sell of strictly positive quantity. Limit orders will fill the book at this price and the expected profit on the $q_{1,t}$th share at price level $p_{1,t}$, is determined by:

$$E[(p_{1,t} - E[v_{t+1}|X_{t+1}] - \gamma)I_{[X_{t+1} > q_{1,t}]}], \tag{3}$$

where $(p_{1,t} - E[v_{t+1}|X_{t+1}])$ is the difference between the price the limit order trades at and the expected fundamental value conditional on the next market order X_{t+1}. $I_{[X_{t+1} > q_{1,t}]}$ is an indicator function that is one if X_{t+1} is larger than $q_{1,t}$—in which case the limit order executes—and zero otherwise. A zero expected profit condition on the last unit (as the queue clears according to first-come-first-served time priority) determines the equilibrium depth $q_{1,t}$ at the best ask price $p_{1,t}$. Once the equilibrium depth is reached, limit order traders will consider submitting a unit at the next price on the grid, one tick above $p_{1,t}$. They consider adding this unit, because the revenue they get on execution is one tick higher than on $p_{1,t}$. The equilibrium depths on both sides of the limit order book are given by the recursions:

$$q_{+k,t} = \frac{p_{+k,t} - v_t - \gamma}{\alpha} - \sum_{i=+1}^{+k-1} q_{i,t} - \lambda \qquad k = 1, 2, \ldots \quad \text{(ask side)} \tag{4}$$

$$q_{-k,t} = \frac{v_t - p_{-k,t} - \gamma}{\alpha} - \sum_{i=-1}^{-k+1} q_{i,t} - \lambda \qquad k = 1, 2, \ldots \quad \text{(bid side)}.$$

The state of the book is described by the set of bid $(p_{-k,t})$ and ask $(p_{+k,t})$ prices and their associated depths $(q_{-k,t}$ and $q_{+k,t})$. Equation 4 shows that the trade informativeness measure

8

α is a key determinant of book liquidity.

3.2 Moment Conditions

We follow Sandås (2001) and use three types of moment conditions: two of these are based on equation (4) where we assume that limit order books have refilled to equilibrium when we take snapshot t (which will be just before the arrival of the next market order X_{t+1}). The third condition identifies the expected market order size λ.

Break-even conditions. For the first type of moment conditions, we use information in the book. In order to eliminate the fundamental value, we add the equilibrium depth associated with the kth price at the bid side of the book from the corresponding equation at the ask side of the book (see equation (4)) and assume that the equations hold up to an error term:

$$E\left(p_{+k,t} - p_{-k,t} - 2\gamma - \alpha\left(\sum_{i=+1}^{+k} q_{i,t} + \sum_{i=-1}^{-k} q_{i,t} + 2\lambda\right)\right) = 0. \qquad k = 1, 2, \ldots \qquad (5)$$

Updating restrictions. For the second type of moment conditions, we use the time dimension and subtract the kth price in the book at time $t - 1$ from the kth price at time t. This removes the fundamental asset value v_t of equation (2), and we get:

$$E\left(\Delta p_{+k,t} - \alpha\left(\sum_{i=+1}^{+k} q_{i,t+1} - \sum_{i=+1}^{+k} q_{i,t}\right) - c - \alpha X_t\right) = 0 \qquad k = 1, 2, \ldots \qquad (6)$$

$$E\left(\Delta p_{-k,t} + \alpha\left(\sum_{i=-1}^{-k} q_{i,t+1} - \sum_{i=-1}^{-k} q_{i,t}\right) - c - \alpha X_t\right) = 0 \qquad k = 1, 2, \ldots$$

where $\Delta p_{k,t} = p_{k,t} - p_{k,t-1}$.

Market order size condition. We use the expected size of the market order to identify λ (see equation (1))via

$$E(|X_t| - \lambda) = 0. \qquad (7)$$

9

3.3 GMM Estimation Results

We use GMM to estimate the structural model based on the four best quotes on both sides of the book. This yields thirteen moment conditions: four from the break-even conditions of equation (5), eight from the updating restrictions of equation (6), and one from the market order size condition of equation (7). The estimation is based on event time, where the arrival of a market order marks an event. In the implementation, we take the snapshot of the order book just ahead of the arrival of the next market order, whereby we assume that durations between trades—time between market order arrivals—are long enough so that competitive limit order traders have time to refill the book in between trades. We consider this a credible assumption for our sample, given the high limit order activity. As pointed out above, on average 6.1 limit orders are submitted in between two market order arrivals (see Table 1).

[insert Table 2 here]

Table 2 reports the GMM estimates of all parameters in the Glosten-model. We estimate on a stock by stock basis, but report averages per trade activity quartile (where Q1 contains the most actively traded stocks) in order to conserve space. We standardise the informativeness measure α to enable meaningful comparison across stocks:[2]

$$\alpha^G = \alpha \frac{\text{€}50,000}{P^2} \tag{8}$$

where P is the average price for the stock throughout the sample period. We use the superscript G to indicate that it is the Glosten-α, which we will later compare to an alternative measure, the Hasbrouck-α. We interpret α^G, in the context of the Glosten-model, as the relative price impact of a €50,000 market order.

Our findings are similar to those reported by Sandås (2001). First, we find that trade informativeness α^G decreases monotonically with trade activity in the cross-section. Market orders seem to be most informative for the least actively traded stocks' quartile (Q4), which contains the smallest stocks in terms of market capitalisation (see Table 1). Second, we find that the transaction cost parameter γ is significantly negative, which is worrisome in the context of the Glosten-model. Third, and most discomforting, the GMM $J-$test rejects the model for 29 out of the 30 stocks.

We proceed with separate estimation of the break-even conditions and the updating restrictions to analyse why the model is rejected. We find that the α estimates based on

the break-even conditions remain significantly positive, whereas those based on the updating restrictions turn insignificant for most of the stocks. It seems that α is identified primarily through snapshots of the book at a single point in time and *not* on updating of the value v_t after the arrival of market orders. This evidence is consistent with Sandås' paper who finds that the Glosten-αs estimated on the break-even conditions are about nine times higher than those based on the updating restrictions (Sandås (2001, Table 5)).

In the second part of our empirical analysis, we deviate from Sandås' (2001) structural approach, and rely on time series econometrics to study the importance of trade informativeness for book liquidity. We abandon the structural model for two reasons. First, the approach requires strong model assumptions in building the moment conditions (e.g. exponential order size distribution, independence of X_t). These assumptions might not be justifiable in real-world markets. Second, the model restricts the same parameter (α) to capture the shape of the book *and* the value update based on arriving market orders. The assumption is that the book replenishes and fully reveals the long-term impact (i.e. the informational content) of the trade before the next market order arrives. In the time series approach, we separate book liquidity from informativeness of market orders to check whether, as the Glosten-model predicts, book liquidity is low during times of highly informative market orders. We rely on Granger causality to identify such effect.

4 Time Series Econometrics: Trade Informativeness and Book Liquidity

In this section, we use the time series dimension to study trade informativeness and order book liquidity. First, we analyse the time series properties of the trade informativeness measure proposed by Hasbrouck (1991). Second, we motivate three measures for book liquidity to capture both the spread and the depth of the book. Finally, we analyse the interaction of trade informativeness and book liquidity based on a Granger-causality test including appropriate control variables to isolate the effects of interest.

For our analysis, we aggregate the event-time calendar used for GMM estimation. In particular, we group N subsequent market orders into one time interval t. N is chosen in order to retain sufficient observations to estimate Hasbrouck's trade informativeness measure. Our procedure is best explained by the following graph where trade arrivals are denoted by

circles and $N = 3$:
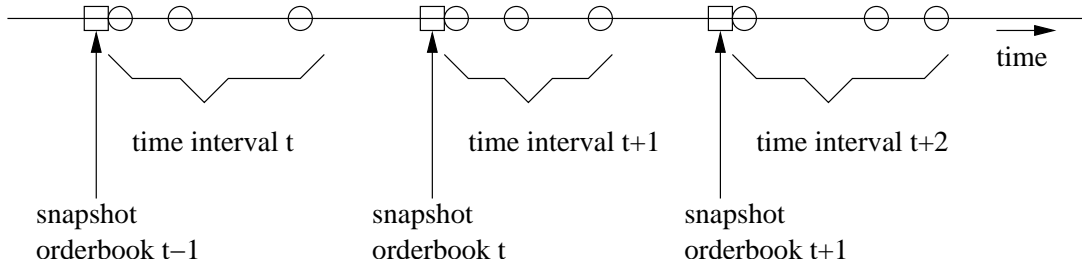
We choose N to trade off too few observations and too little time variation. We therefore set N equal to 250 for the quartile of most actively traded stocks (Q1) yielding an average of 14.9 periods per trading day, each lasting 34 minutes on average. We take less observations per interval for the three other quartiles by setting $N = 200$ for Q2, $N = 150$ for Q3, and $N = 100$ for Q4. By definition, these quartiles trade less frequently than Q1 so that for equal N we would obtain too few intra-day estimation intervals. Also, these smaller values of N ensure that the average time lengths of the intervals are comparable to Q1. The resulting average interval lengths are 43 minutes for Q2, 61 minutes for Q3, and 60 minutes for Q4. For robustness checks, we try various values for N and find that our main results are not affected.

4.1 Measuring Time-Varying Trade Informativeness

We use the N trades in a time interval to identify trade informativeness following closely Hasbrouck (1991). The approach relies on estimates of the following bivariate vector autoregression (VAR):

$$
\begin{aligned}
r_{t,\tau} &= a_{1t}r_{t,\tau-1} + a_{2t}r_{t,\tau-2} + \ldots + b_{0t}X_{t,\tau} + b_{1t}X_{t,\tau-1} + b_{2t}X_{t,\tau-2} + \ldots + u_{t,\tau} \quad (9)\\
X_{t,\tau} &= c_{1t}r_{t,\tau-1} + c_{2t}r_{t,\tau-2} + \ldots + d_{1t}X_{t,\tau-1} + d_{2t}X_{t,\tau-2} + \ldots + w_{t,\tau}
\end{aligned}
$$

where $\tau \in \{1, \ldots, N\}$ runs over all midquote changes in time interval t, $r_{t,\tau}$ is the midquote change from order book snapshot $\tau - 1$ to snapshot τ in interval t, and $X_{t,\tau}$ is the signed order size of trade τ in interval t (it is zero if there was no transaction at the time of the midquote change). We use the subscript t to indicate the coefficients that belong to time interval t. We assume that the innovations $u_{t,\tau}$ and $w_{t,\tau}$ are i.i.d. and contemporaneously

uncorrelated.

To identify the long-term impact of a trade, we consider the vector moving average representation of the VAR in equation (9) (assuming stationarity):

$$r_{t,\tau} = u_{t,\tau} + a_{1t}^* u_{t,\tau-1} + a_{2t}^* u_{t,\tau-2} + \ldots + b_{0t}^* w_{t,\tau} + b_{1t}^* w_{t,\tau-1} + \ldots \tag{10}$$

$$X_{t,\tau} = c_{1t}^* u_{t,\tau-1} + c_{2t}^* u_{t,\tau-2} + \ldots + w_{t,\tau} + d_{1t}^* w_{t,\tau-1} + \ldots \tag{11}$$

where b_{0t}^* identifies the immediate price impact of a market order. The permanent price impact of a market order in time interval t can now be obtained as:

$$\alpha_t^H = (\sum_{i=0}^{\infty} b_{it}^*)\frac{\text{€}50,000}{P^2} \tag{12}$$

where the second factor on the right-hand side of the equation scales the coefficient in the same way as was done for the Glosten-α (see equation (8)).

[insert Table 3 here]

In the VAR estimations, we use ten lags to ensure uncorrelated residuals. Hasbrouck (1991) chooses five VAR lags, and Dufour and Engle (2000), who extend Hasbrouck's methodology, choose the same lag length. The choice of ten lags in our sample ensures serially uncorrelated VAR residual series for all stocks. The variation of the lag length within the range indicated by Akaike information criteria does not affect the results of the later stages which make use of the trade informativeness estimates. The increased lag length which is indicated for our recent data supports the view that order splitting/algorithmic trading became more important since the early 1990s.

We follow Hasbrouck (1991) and truncate the infinite sum of equation (12) at lag forty. As above, we estimate the VAR stock by stock and report the quartile averages.

Table 3 reports the average Hasbrouck-α as well as its time series characteristics. We find that the average Hasbrouck-α is similar to the average Glosten-α. For the quartile of most actively traded stocks (Q1), we find an average α^H of 0.08 basis points, which is very close to the average α^G of 0.09 basis points reported in Table 2. For the other quartiles, the differences are somewhat larger, and the Hasbrouck-α is consistently smaller than the Glosten-α. We interpret the similarity of the Glosten- and Habsrouck-α estimates, which are

13

obtained by very different methodologies, as further evidence that the updating restrictions are a poor measure of trade informativeness in the GMM test of the Glosten-model. In Section 3.3, we showed that separate estimation of the break-even conditions (based on book depth) and updating restrictions shows that the (overall) Glosten-α estimate largely captures book depth, as it is insignificant in the updating restrictions. Apparently, order book changes in between market orders, captured by these updating restrictions, are not able to identify trade informativeness. The Hasbrouck-α captures the long-term impact of trades and its similarity to the Glosten-α can therefore be interpreted as support for the Glosten-model.

[insert Figure 1 here]

We further find that trade informativeness is a persistent, mean-reverting process with a distinct intra-day pattern. Figure 1 plots the average α^H for 90-minute intervals within the trading day. We generally find that informativeness decreases during the course of the day, which is consistent with the hypothesis that the bulk of price discovery is taking place after the opening of the market. For the quartile of most actively traded stocks (Q1), we find a significant increase of trade informativeness during the interval from 15:00 to 16:30 C.E.T., which contains the NYSE opening time (15:30 C.E.T.). Not surprisingly, all Q1 shares are cross-listed in New York, whereas few of the stocks in the other quartiles are inter-listed. Menkveld (2008) finds the same pattern for British and Dutch ADRs and argues that some traders prefer to trade during the overlap and split orders in order to benefit from two pools of liquidity.

We remove intra-day seasonality (diurnality) of the α^H estimates by dividing by the time-of-day means, and then compute first-order autocorrelations. We find persistence as first-order autocorrelations are significant for all quartiles, ranging from 0.28 to 0.38 (see Table 3). We also report inter- and intra-day correlations separately, and find that both are positive, although only intra-day autocorrelations are statistically significant.

The substantial time variation and predictability in trade informativeness is useful to analyse how important trade informativeness is for book liquidity. The persistence suggests that informative trades cluster in time so that we can discriminate times of informative market order trading and times with relatively uninformed market orders. We will study book liquidity at these times to gauge the importance of trade informativeness. Before we turn to this analysis, we first construct appropriate measures of book liquidity.

14

4.2 Measuring Time-Varying Book Liquidity

We summarise limit order book liquidity at snapshot t (taken at the end of the tth estimation window) through three measures: the quoted bid-ask spread and two measures of book depth. In the remainder of the paper, we focus on the results for the ask side of the limit order book, and leave out the bid side to conserve space. The results are very similar for both sides.

Our depth measures captures the price concession from the best ask rather than the midquote in order to purge the measure of a quoted spread effect (measured separately). Price concessions are calculated for an average-sized and a large buy value. Price concessions are calculated relative to a book state where the depth at the best ask can fill the total volume of the buy order. Formally, we use

$$ap(V) = \frac{\sum_{k=0}^{\infty} I_k(V) p_k q_k}{\sum_{k=0}^{\infty} I_k(V) p_1 q_k} - 1 \tag{13}$$

where

$$I_k(V) = \begin{cases} 1 & \text{if } \sum_{i=1}^{k} q_i p_i \leq V \\ \frac{V - \sum_{i=1}^{k-1} q_i p_i}{q_k p_k} & \text{if } \sum_{i=0}^{k-1} q_i p_i \leq V < \sum_{i=1}^{k} q_i p_i \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

For ease of notation, we define a summation from i to j with $j < i$ to be zero (as this happens in the second line in equation (14) for $k = 1$. The depth measure $ap(V)$ is zero if the market buy order of size V fills at the best ask and becomes strictly positive if it has to walk up the book and consume the depth displayed at higher prices. A large price concession indicated by a large $ap(V)$ therefore indicates poor book liquidity at and behind the best quote. We compute the depth measure (13) for $V = $ €50,000 and €200,000, respectively. In the remainder of the paper, we use $ap(50)$ and $ap(200)$, respectively, to refer to these depth measures.

4.3 Informativeness and Liquidity: Granger Causality

We use time-varying trade informativeness and book liquidity to determine whether trade informativeness is an important determinant of book liquidity. We benefit from slow mean-reversion in the trade informativeness process to discriminate highly informative from rel-

atively uninformative trade periods. We also study whether book liquidity predicts next period trade informativeness, which could be the case when limit order submitters know of an oncoming information event that is not predicted by the informativeness of lagged order flow (e.g. a pre-scheduled company news release).

For the remainder of the analysis we standardise all data series—α and the book liquidity measures—in order to be able to compare across intra-day time intervals and across stocks. For that purpose, we divide the trading day into six 90-minute time intervals (see also Figure 1). For each stock and each variable, we then remove the diurnality by dividing the time series x_t by its time-of-day mean (i.e. we replace x_t by $x_t/\overline{x}_i$ where $i \in \{1, \ldots, 6\}$ is the time interval t falls into and $\overline{x}_i$ is the average of all x_t in that interval). Furthermore, in order to account for heteroskedasticity across stocks, we scale the resulting time series by their stock-specific standard deviation (i.e. we replace the diurnally adjusted x_t by x_t/σ_{x_t} where σ_{x_t} is the standard deviation of the diurnally adjusted x_t). The slope coefficients in the Granger-causality analysis below should therefore be interpreted as the amount of change of the dependent variable (in terms of its standard deviation) on a one standard deviation change of the explanatory variable.

4.3.1   Trade informativeness as a determinant of book liquidity.

Before turning to the Granger causality analysis, we first plot book liquidity against trade informativeness. By construction, the book snapshot t is at the end of time interval t, which contains the N trades that are used to calculate α^H (see graph at the start of Section 4). To study whether informativeness matters for subsequent book liquidity, we therefore plot the three book liquidity measures against α_t^H where we group the α into quartiles.

[insert Figure 2 here]

Figure 2 illustrates that book liquidity decreases monotonically in trade informativeness. Panel A illustrates that the bid-ask spread increases linearly in trade informativeness. The spread is 10% higher at times of highest trade informativeness relative to times of lowest trade informativeness. Panel B and C illustrate that trade informativeness has an even stronger effect on the two depth measures ($ap(50)$ and $ap(200)$, respectively) as depth is 20% lower comparing the two tail quartiles. And, the marginal effect seems to be increasing in trade informativeness.

16

We now turn to Granger causality regressions to study whether trade informativeness *causes* book liquidity. We propose the following regression:

$$y_t = c + \delta \alpha_t^H + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p} + \phi_1 z_{1,t} + \ldots + \phi_q z_{q,t} + \varepsilon_t \qquad (15)$$

where y_t is one of the three book liquidity measures (*spread*, $ap(50)$, or $ap(200)$), $z_{i,t}$ are control variables that are based on the N trades in time interval t to capture market conditions other than trade informativeness. ε_t is an i.i.d. disturbance. The regression indicates Granger causality if δ is statistically significant so that trade informativeness explains book liquidity over and above its predicted value based on its own past. By adding control variables, effectively, we consider the component of trade informativeness that is orthogonal to other indicators measuring trading conditions. If, in this case, δ is significant, we can attribute the effect uniquely to informativeness and not to other correlated trading variables.

[insert Table 4 here]

Table 4 shows all control variables, their pairwise correlations, and their correlations with α_t^H and the three book liquidity measures. We use evidence from the microstructure literature (see Madhavan (2000) or Biais et al. (2005) for a survey) to motivate five control variables. First, we include the trade size averaged over the estimation interval (*tsize*). This accounts for a potential size effect driving the significance of the trade informativeness measure if large trades are more informative than small trades. Second, we include the signed trade volume imbalance over the estimation interval (*simb*), in order to account for an asymmetric effect on book liquidity. Third, we add mean inter-trade time duration over the estimation interval (*dura*) to control for fast markets that allegedly indicate more informative trades (see Dufour and Engle (2000)). Fourth, we compute *realised volatility* (*rv*) over the estimation interval. Including it as another control variable we aim to differentiate trade informativeness from picking-off risk induced by volatility shocks caused by the arrival of public information. *rv* is computed by summing squared midquote returns within the estimation interval. Midquotes are sampled immediately before the trade events within the interval. The idea to use high frequency sampled midquote changes to estimate and forecast lower frequency fundamental price volatility goes back to Andersen et al. (2003). Fifth, we add the absolute value of trade imbalance (*imb*) to control for trading volume. We prefer this proxy to the conventional volume measure as it better captures the net pressure on liquidity suppliers. The correlations of these five control variables with α_t^H are significantly positive for trade size, duration, realised volatility, and volume. Periods with

17

highly informative trades, therefore, tend to show large trades in slow markets with high volatility and high volume. These correlations are as expected, except maybe for inter-trade duration. Short trade durations in the Xetra limit order book setting are associated with relatively uninformative trades. This is in contrast to the findings reported in Dufour and Engle (2000) for the 1991 NYSE market. However, Grammig et al. (2007), who propose a structural modelling alternative to Dufour and Engle's (2000) time series approach, also find that fast trading in the Xetra limit order book cannot be associated with informed trading. They refer to the *crowding-out* effect described by Parlour (1998) to explain their finding. Crowding-out means that ample liquidity - a market state that we do not associate with asymmetric information - causes intense uninformed market order trading. By submitting a market order, an impatient (yet not superiorly informed) trader can jump ahead of a lengthy queue of limit orders at the best quotes.

[insert Table 5 here]

Panel A of Table 5 summarises the regression results using the quoted spread as dependent variable. Trade informativeness is significant for eight out of thirty stock-specific regressions. The parameter carries the correct sign, since at times of high trade informativeness the book exhibits higher bid-ask spreads. We find that the effect is economically significant as a one standard deviation change of trade informativeness increases the spread by 8% of its standard deviation. The quartile-specific results show that the effect increases monotonically in trade activity. It is strongest for the least actively traded stocks. Regarding the control variables, we find that realised volatility is the only variable more often significant than trade informativeness (for fourteen stocks). It also has a stronger economic significance, as a one standard deviation increase in realised volatility increases the spread by 15% of its standard deviation.

Panels B and C of Table 5 report the results for the two depth measures $ap(50)$ and $ap(200)$ which show that trade informativeness is important primarily for depth further into the book. Using $ap(50)$ as dependent variable in (15), trade informativeness is statistically significant for six stocks, while it is statistically significant for seventeen sample stocks when $ap(200)$ is used as the dependent variable in (15). The estimated coefficients have the correct sign, as high informativeness implies higher price concessions. The effects are economically significant as a one standard deviation change in trade informativeness increases $ap(50)$ by 7% and $ap(200)$ by 12% of the respective standard deviation. Trade informativeness is economically and statistically more important than realised volatility, which is the second-most relevant variable. For $ap(200)$, we again find a monotonic relationship across activity

18

quartiles, where the least active quartile shows the strongest effect.

4.3.2 Book liquidity as a determinant of trade informativeness.

The Glosten model does not only predict that informativeness Granger-causes book liquidity, but also implies reverse causality. If limit order traders who fill the book at the end of period t know more about the oncoming period than what could be predicted from current and past period informativeness, the book could Granger-cause next period informativeness. For example, traders might know that a company is about to issue a press release, which, of course, leads to an immediate quote update, but also to market orders that are highly informative. This is the case when the public information leads to "allocational" trades due to portfolio re-balancing based on public news (see Vayanos (2001)). The econometrician cannot predict such event based on past trade informativeness.

We test whether book liquidity Granger-causes trade informativeness through the following regression:

$$\alpha_t^H = c + \beta_1 \alpha_{t-1}^H + \ldots + \beta_p \alpha_{t-p}^H + \delta_1 y_{1,t-1} + \ldots + \delta_r y_{r,t-1} + \phi_1 z_{1,t-1} + \ldots + \phi_q z_{q,t-1} + \varepsilon_t \quad (16)$$

where the control variables $z_{i,t}$ are the same as in equation (15). For the measures of book liquidity, $y_{i,t}$, we use quoted spread and our two price concession measures $ap(50)$ and $ap(200)$. We add a book asymmetry measure that is defined as the ask price concession minus the bid price concession: $dpi50 = ap(50)$-$bp(50)$ and $dpi200 = ap(200)$-$bp(200)$. A low value indicates that it is more expensive to sell than to buy, which might foreshadow a downturn. Facing a downturn threat, limit order traders might rely more on the direction of future market orders for price discovery. This would imply increased trade informativeness.

[insert Table 6 here]

The regression results in Table 6 indicate that book liquidity Granger-causes trade informativeness. We find that a high quoted spread causes trades to be significantly more informative in the oncoming period for 11 out of 30 stocks. In addition, we find that low depth causes increased trade informativeness, albeit primarily for $ap(200)$. We find it to be significant for 22 stocks. Interestingly, the book depth asymmetry measure for large trades is also significant for 15 stocks. The negative sign of the parameter estimate is consistent with the intuition that market orders are more informative when selling is relatively expensive. As

19

for the control variables, we find strong statistical significance only for trade size, as large trades predict lower trade informativeness in the oncoming period for 17 stocks.

5   Discussion of the Results

Overall, we find evidence that trade informativeness matters for book liquidity in limit order markets. Along with realised volatility it performs better than any other variable that captures trading conditions such as trading intensity, trading volume and order book imbalance. There are, however, some issues that are worthy of discussion.

First, we want to emphasise that book liquidity and trade informativeness are not the same thing. It is true that on poor book liquidity subsequent market orders have a large price concession by definition. However, it is not the immediate price impact that represents the adverse-selection that the Glosten model accounts for. In fact, part of the price concession is temporary and exists to compensate liquidity suppliers for their order-processing cost. In theory, if market orders are not informative (as assumed in most of the dynamic models) all of the price concession would be temporary.

Second, the Granger causality results warrant some discussion. When constructing the data series for the causality test, we are forced to aggregate across market orders and create time intervals. We do this to estimate the Hasbrouck model and determine the long-term price impact of a trade. In the test, we pursue the Granger causality idea that one should have explanatory power beyond from what can be predicted from a series' own past. One could argue that we have not accurately controlled for book liquidity's own past, as we include only lagged book snapshots from before the trade interval. Ideally, we want to include the snapshot from before the previous market order, but we cannot do this as we need the aggregation to identify the *long-term* price impact. An alternative interpretation of our test is that it does show that trade informativeness is significant in explaining subsequent book liquidity beyond the prediction from its own past from *before* the trade interval.

6   Conclusion

We analyse three months of limit order book data—January through March 2004—from the German Stock Exchange for the thirty DAX stocks to test the predictions of the Glosten (1994) model. One of the key predictions is that order books are filled by competitive limit order traders, so that, in equilibrium, the marginal order just earns enough to make up for order-processing cost and the adverse-selection cost of executing against a (potentially) privately informed market order.

First, we follow Sandås (2001) and estimate the parameters of the Glosten model directly using GMM. We reject the model econometrically and diagnose that the key parameter α, which measures the level of private information in market orders, is primarily identified on book restrictions and *not* on the updating restrictions that should track the information in the trade. Second, we leave the structural model and exploit the time dimension to study whether trade informativeness matters for book liquidity. We develop measures for book liquidity and use the Hasbrouck (1991) methodology to gauge trade informativeness. We document that trade informativeness Granger-causes book liquidity with the strongest statistical and economical significance for large order price concessions ("behind the market"). To robustify these results we control for various other determinants of LOB liquidity, such as realised volatility, trading intensity, trade size, and order book imbalance. Among these, we find that only realised volatility rivals trade informativeness in terms of economic and statistical significance. In particular, realised volatility is more relevant for top-of-the-book liquidity, whereas informativeness is more important for beyond-the-best-quote liquidity.

We interpret these results as support for prominent theoretical models of limit order book markets. First, our finding that trade informativeness is one of the most important explanatory factors for book liquidity supports Glosten's (1994) model which explains order book equilibrium with order-processing costs and information asymmetry. Second, the finding that realised volatility is more important than trade informativeness for top-of-the-book liquidity supports a key result of dynamic limit order book models like that developed by Foucault (1999). These models start from symmetric information and assume that the key cost to limit orders is picking-off risk. That is, limit orders are costly if they are consumed before cancelled on the arrival of public information. This cost is particularly relevant "at the market", which explains the result that realised volatility is more important than trade informativeness for top-of-the-book liquidity.

Notes

[1]We define "price concession" as the hypothetical transaction price for a given trade volume relative to a reference price (e.g., the midquote or the best quote). For limit order markets, we prefer price concession to the bid-ask half spread as a measure of liquidity, since depth at the best quote is often too small to transact the market order, which then runs up the book.

[2]Traders seem to transact in terms of value, not in terms of number of shares. That is, we find that the average value per trades is similar across stocks, which is not true for the average number of shares per trade. The reason is that, in the cross-section, stocks trade at different price levels.

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(2), 767–788.

ANDERSEN, T., T. BOLLERSLEV, F. DIEBOLD, AND P. LABYS (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

BIAIS, B., T. FOUCAULT, AND F. SALANIÉ (1998): "Floors, Dealer Markets, and Limit Order Markets," *Journal of Financial Markets*, 1, 253–284.

BIAIS, B., L. GLOSTEN, AND C. SPATT (2005): "A Survey of Microfoundations, Empirical Results, and Policy Implications," *Journal of Financial Markets*, 8, 217–264.

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.

BIAIS, B., D. MARTIMORT, AND J. ROCHET (2000): "Competing Mechanisms in a Common Value Environment," *Econometrica*, 68, 799–837.

BLOOMFIELD, R., M. O'HARA, AND G. SAAR (2009): "How Noise Trading Affects Markets: An Experimental Analysis," *Review of Financial Studies*, 22(6), 2275–2302.

DUFOUR, A., AND R. ENGLE (2000): "Time and the Price Impact of a Trade," *Journal of Finance*, 55(6), 2467–2499.

EASLEY, D., S. HVIDKJAER, AND M. O'HARA (2002): "Is Information Risk a Determinant of Asset Returns?," *Journal of Finance*, 57, 2185–2221.

EASLEY, D., N. KIEFER, M. O'HARA, AND J. PAPERMAN (1996): "Liquidity, Information, and Less-Frequently Traded Stocks," *Journal of Finance*, 51, 1405–1436.

FOUCAULT, T. (1999): "Order Flow Composition and Trading Costs in a Dynamic Limit Order Market," *Journal of Financial Markets*, 2, 99–134.

FOUCAULT, T., O. KADAN, AND E. KANDEL (2005): "Limit Order Book as a Market for Liquidity," *Review of Financial Studies*, 18(4), 1171–1217.

GLOSTEN, L. (1994): "Is the Electronic Limit Order Book Inevitable?," *Journal of Finance*, 49, 1127–1161.

GOETTLER, R., C. PARLOUR, AND U. RAJAN (2005): "Equilibrium in a Dynamic Limit Order Market," *Journal of Finance*, 60(5), 2149–2192.

——— (2009): "Informed Traders and Limit Order Markets," *Journal of Financial Economics*, 93, 67–87.

GRAMMIG, J., E. THEISSEN, AND O. WÜNSCHE (2007): "Time and Price Impact of a Trade: A Structural Approach," Discussion paper, EFA 2007 Conference Paper, Universities Tübingen and Mannheim.

GRIFFITHS, M., B. SMITH, D. TURNBULL, AND R. WHITE (2000): "The Costs and Determinants of Order Aggressiveness," *Journal of Financial Economics*, 56, 65–88.

HARRIS, L., AND J. HASBROUCK (1996): "Market vs. Limit Orders: the SuperDOT Evidence on Order Submission Strategy," *Journal of Financial and Quantitative Analysis*, 31, 213–231.

HASBROUCK, J. (1991): "Measuring the Information Content of Stock Trades," *Journal of Finance*, 46, 179–207.

HASBROUCK, J., AND G. SAAR (2007): "Limit Orders and Volatility in a Hybrid Market: The Island ECN," Discussion paper, New York University.

——— (2009): "Technology and Liquidity Provision: The Blurring of Traditional Definitions," *Journal of Financial Markets*, 12(2), 143–172.

HOLLIFIELD, B., R. MILLER, AND P. SANDÅS (2004): "Empirical Analysis of Limit Order Markets," *Review of Economic Studies*, 71, 1027–1063.

——— (2006): "Estimating the Gains from Trade in Limit Order Markets," *Journal of Finance (forthcoming)*.

HUANG, R., AND H. STOLL (1997): "The Components of the Bid-Ask Spread: A General Approach," *Review of Financial Studies*, 10(4), 995–1034.

MADHAVAN, A. (2000): "Market Microstructure: A Survey," *Journal of Financial Markets*, 3(3), 205–258.

MENKVELD, A. (2008): "Splitting Orders in Overlapping Markets: A Study of Cross-Listed Stocks," *Journal of Financial Intermediation*, 17, 145–174.

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

PARLOUR, C., AND D. SEPPI (2003): "Liquidity-Based Competition for Order Flow," *Review of Financial Studies*, 16(2), 301–343.

RANALDO, A. (2004): "Order Aggressiveness in Limit Order Book Markets," *Journal of Financial Markets*, 7, 53–74.

ROSU, I. (2009): "A Dynamic Model of the Limit Order Book," *Review of Financial Studies*, 22, 4601–4641.

SANDÅS, P. (2001): "Adverse Selection and Competitive Market Making: Empirical Evidence from a Limit Order Market," *Review of Financial Studies*, 14, 705–734.

SEPPI, D. (1997): "Liquidity Provision with Limit Orders and a Strategic Specialist," *Review of Financial Studies*, 10, 103–150.

VAYANOS, D. (2001): "Strategic Trading in a Dynamic Noisy Market," *Journal of Finance*, 56, 131–171.

Table 1: Characteristics of the stocks in the sample (DAX30 stocks).

Ticker	Company Name	market cap. (mill. €)	% agg. trades	daily nb. trades	daily nb. subm.	daily nb. cancel.	price (€)	spread (€)	spread (%)	activity quartile
ALV	ALLIANZ	33,805	21.4	4,523	29,791	25,882	100.1	0.05	0.05	
DTE	DEUTSCHE TELEKOM	34,858	5.0	4,445	14,498	11,009	15.7	0.01	0.07	
SIE	SIEMENS	52,893	16.7	4,418	23,659	19,920	64.0	0.03	0.05	Q1
DBK	DEUTSCHE BANK	38,228	19.3	3,961	23,169	19,772	67.2	0.03	0.05	
MUV2	MUENCHENER RUECK	16,396	20.7	3,425	20,154	16,894	93.9	0.06	0.06	
DCX	DAIMLERCHRYSLER	30,316	14.5	3,309	18,722	15,919	36.4	0.02	0.06	
EOA	E.ON	33,753	13.6	2,871	18,899	16,468	52.5	0.03	0.06	
SAP	SAP	27,412	21.9	2,806	19,733	17,095	131.5	0.08	0.06	
IFX	INFINEON	4,790	8.6	2,799	10,320	7,744	11.6	0.01	0.10	
BAS	BASF	25,425	13.8	2,580	18,211	15,898	43.3	0.03	0.06	Q2
VOW	VOLKSWAGEN	9,688	16.0	2,545	13,474	11,273	39.2	0.03	0.07	
BAY	BAYER	15,911	12.4	2,400	15,258	12,988	23.1	0.02	0.08	
RWE	RWE	12,653	13.0	2,314	14,438	12,355	33.8	0.03	0.08	
BMW	BMW	12,211	14.4	2,110	14,736	12,764	34.7	0.02	0.07	
HVM	BAY.HYP.VEREINSBANK	6,629	15.0	1,937	10,204	8,293	18.7	0.02	0.11	
SCH	SCHERING	7,055	16.2	1,523	9,111	7,669	40.8	0.04	0.09	
CBK	COMMERZBANK	7,569	12.6	1,450	11,922	10,476	15.4	0.02	0.11	
LHA	LUFTHANSA	4,548	11.9	1,352	8,079	6,780	14.2	0.02	0.12	Q3
DPW	DEUTSCHE POST	6,806	11.0	1,315	6,861	5,666	18.2	0.02	0.11	
TKA	THYSSEN-KRUPP	6,450	11.3	1,262	7,864	6,672	15.9	0.02	0.13	
MEO	METRO	5,018	15.7	1,235	7,975	6,702	35.0	0.04	0.12	
ALT	ALTANA	3,338	18.9	1,095	7,718	6,609	48.6	0.05	0.10	
TUI	TUI	2,025	17.6	1,063	6,767	5,714	18.7	0.03	0.14	
MAN	MAN	2,434	13.0	1,057	7,214	6,235	27.7	0.03	0.12	
CONT	CONTINENTAL	4,060	13.5	1,002	8,036	7,052	31.6	0.04	0.11	
DB1	DEUTSCHE BOERSE	4,847	18.4	982	6,598	5,698	46.9	0.04	0.10	Q4
ADS	ADIDAS-SALOMON	4,104	20.1	980	8,057	7,105	92.6	0.08	0.09	
LIN	LINDE	3,448	15.8	896	8,342	7,454	43.6	0.05	0.11	
HEN3	HENKEL	3,682	16.6	702	7,989	7,306	65.9	0.07	0.10	
FME	FRESENIUS	1,944	16.7	621	5,764	5,195	54.0	0.07	0.13	
	Average	14,076	15.2	2,099	12,785	10,887	44.5	0.04	0.09	

26

Note: The statistics are computed based on market event data covering the sample period January 2, 2004 to March 31, 2004. Column *market cap.* gives the market capitalisation of the respective stock in million euros at the end of December 2003. *% agg. trades* is the percentage of total trading volume executed beyond the best quotes (aggressive trades). *daily nb. trades* denotes the average daily number of trades. Column *daily nb. subm.* reports the average number of order submissions per day, market orders excluded and *daily nb. cancel.* the average number of order cancellations per day. *price* (€), *spread* (€) and *spread (%)* are average midquote, spread and relative spread over the three months sample period. The stocks are sorted into four groups (activity quartiles, Q1-Q4) according to their trading frequency, i.e. by the column *daily nb. trades*. Horizontal lines separate the four groups.

Table 2: Estimation results of the Såndas/Glosten model

.

	$\alpha \times 10^3$	γ	λ	$c_v \times 10^3$	$J(9)$	no reject	$\alpha^G \times 10^3$
all stocks	**0.013**	**-0.010**	**1.380**	**0.030**	643.7	2	**0.326**
	167.3	52.5	133.1	2.2			
Q1 (most active)	**0.009**	**-0.006**	**1.600**	**0.029**	1985.3	0	**0.087**
	249.6	50.5	188.8	2.5			
Q2	**0.008**	**-0.008**	**1.724**	**0.088**	539.8	0	**0.189**
	170.3	55.7	147.1	2.0			
Q3	**0.008**	**-0.010**	**1.517**	**-0.063**	91.9	1	**0.471**
	143.1	54.9	108.5	2.1			
Q4 (least active)	**0.029**	**-0.018**	**0.610**	**0.071**	51.6	1	**0.555**
	109.3	48.2	89.4	2.1			

Note: The estimation uses the information from the best four quotes of the visible books to form update and break-even conditions. The table reports in bold face font the first stage GMM estimates of α, γ, λ and c_v which are averaged across all stocks as well as across the stocks in the respective trading activity quartile. The values printed in regular font are $t-$values which are also group averages. The $J(9)$ column is the group average of the GMM $J-$ statistic (with 9 degrees of freedom). Column *no reject* reports the number of stocks for which the model is not rejected at 1% significance level. The last column reports the group averages of the standardised trade informativeness measure α^G (Glosten-α, see equation (8)).

Table 3: Cross-sectional and time series properties of estimated trade informativeness measures α^H and immediate impacts b^H.

	$\alpha^H \times 10^3$	min	max	ρ	ρ_{inter}	ρ_{intra}	$b^H \times 10^3$
all stocks	**0.244**	**-0.016**	**0.963**	**0.328**	**0.162**	**0.352**	**0.139**
Q1	**0.076**	**0.005**	**0.279**	**0.365**	**0.158**	**0.387**	**0.042**
(most active)	0.021	0.002	0.056	0.051	0.166	0.055	0.011
Q2	**0.152**	**-0.011**	**0.649**	**0.355**	**0.168**	**0.380**	**0.089**
	0.041	0.034	0.260	0.114	0.181	0.113	0.025
Q3	**0.335**	**-0.006**	**1.282**	**0.282**	**0.155**	**0.309**	**0.199**
	0.094	0.015	0.386	0.067	0.185	0.061	0.056
Q4	**0.412**	**-0.055**	**1.639**	**0.311**	**0.166**	**0.333**	**0.224**
(least active)	0.122	0.066	0.422	0.046	0.088	0.058	0.074

Note: The table reports sample means (bold font) and standard deviations (regular font) of α^H (Hasbrouck-α (see equation (12)) and the immediate impact computed as $b^H = \frac{b_0 50,000}{P^2}$ (see equations 10-12). Both α^H and b^H are diurnally adjusted. For that purpose, the trading day is divided into 90 minutes intervals and the (stock specific) sample mean of α^H and b^H in each interval is computed. Diurnally adjusted variables result from dividing the raw series by the corresponding time-of-day means. To compute sample means and standard deviations reported in the table, the diurnally adjusted series are pooled (overall and by trade size quartile, respectively). The columns labelled min and max report the group averages of the smallest and largest α^H estimate ($\times 10^3$) for each stock. Column ρ reports the autocorrelation of the diurnally adjusted α^H sequence. ρ_{inter} is the inter-day correlation of the diurnally adjusted Hasbrouck-α, i.e. the correlation between the previous day's last and the next day's first α^H estimate. ρ_{intra} is the intra-day serial correlation of the diurnally adjusted α^H. Observations from different trading days are excluded for this computation. Bold (regular) faced numbers are group averages (standard deviations) of the autocorrelations computed in this way.

Table 4: Correlation between the trade informativeness, liquidity measures and control variables.

	α^H	$ap(50)$	$ap(200)$	$spread$	$tsize$	$simb$	$dura$	rv
$ap(50)$	0.11**							
$ap(200)$	0.19**	0.69**						
$spread$	0.11**	-0.04**	0.00					
$tsize$	-0.46**	-0.06**	-0.11**	-0.01				
$simb$	-0.02*	0.00	-0.01	0.02*	0.03**			
$dura$	0.11**	0.01	-0.01	-0.02**	-0.34**	-0.01		
rv	0.28**	0.08**	0.14**	0.17**	0.08**	0.00	-0.14**	
imb	-0.21**	-0.03**	-0.04**	0.04**	0.47**	0.06**	-0.22**	0.06**

Note: The table reports cross sectional averages of correlation coefficients for the DAX30 stocks. $ap(50)$ and $ap(200)$ denote the hypothetical ask side price concessions for trades of €50,000 and €200,000, respectively. $spread$ is the quoted spread in basis points. These liquidity measures are sampled just before the first trade occurs after the α^H estimation interval. $tsize$ denotes the trade size and $dura$ the duration between two consecutive trades averaged over the estimation interval. $simb$ is the sum of the signed traded volumes over the estimation interval; imb is the absolute value of $simb$. Realised volatility (rv) is computed as the sum of the squared midquote returns with midquotes sampled just before a trade event occurs within the estimation interval. All series are diurnally adjusted. For that purpose the trading day is divided into 90 minutes intervals and the (stock specific) mean of the respective variable in each interval is computed. Diurnally adjusted variables result from dividing the raw series by their corresponding time-of-day means. Computation of the correlations is based on an average (across stocks) of $T = 692$ observations. ** and * indicate a mean correlation that is significantly different from zero at the 1 % and 5 % level, respectively. The standard errors of the mean correlations are approximated by $\frac{1}{\sqrt{TN}}$ where $N = 30$.

Table 5: Regression of book liquidity measures on trade informativeness and control variables.

		α^H	rv	$tsize$	$dura$	imb	$simb$	$lag1$	$lag2$	R^2_{adj}
Panel A *spread*	All stocks	0.080	0.147	-0.015	-0.005	0.047	0.014	0.001	0.004	0.051
	sig_pos (of 30)	8	14	0	2	5	0	0	0	
	sig_neg	0	0	3	3	0	1	1	0	
	Q1 (most active)	0.043	0.125	-0.065	-0.037	0.035	0.034	-0.010	0.003	0.034
	Q2	0.068	0.168	-0.012	-0.015	0.043	0.007	-0.003	-0.012	0.061
	Q3	0.089	0.192	0.002	0.005	0.036	-0.007	-0.007	0.024	0.067
	Q4 (least active)	0.121	0.095	0.011	0.025	0.076	0.024	0.024	0.002	0.038
Panel B *ap(50)*	All stocks	0.068	0.063	-0.041	-0.008	0.004	0.003	0.009	0.005	0.015
	sig_pos (of 30)	6	5	0	0	0	1	0	0	
	sig_neg	0	0	4	0	1	1	1	1	
	Q1 (most active)	0.077	0.047	0.022	-0.007	-0.030	0.018	0.002	0.012	0.010
	Q2	0.060	0.041	-0.069	-0.010	0.034	0.013	0.014	-0.015	0.010
	Q3	0.043	0.110	-0.056	-0.017	-0.009	-0.022	-0.001	0.013	0.019
	Q4 (least active)	0.097	0.050	-0.055	0.004	0.017	0.006	0.021	0.011	0.022
Panel C *ap(200)*	All stocks	0.117	0.096	-0.072	-0.030	0.014	-0.004	0.041	0.042	0.054
	sig_pos (of 30)	17	13	0	0	1	1	1	4	
	sig_neg	0	0	5	1	1	2	0	1	
	Q1 (most active)	0.108	0.099	-0.035	-0.036	0.005	0.001	-0.010	0.018	0.034
	Q2	0.107	0.085	-0.082	-0.034	0.041	-0.006	0.052	0.030	0.048
	Q3	0.124	0.126	-0.073	-0.030	-0.004	-0.005	0.046	0.046	0.064
	Q4 (least active)	0.129	0.071	-0.096	-0.018	0.012	-0.006	0.076	0.076	0.071

Note: The book liquidity indicators *spread*, *ap(50)* and *ap(200)* are regressed on the trade-informativeness indicator α^H and control variables. The liquidity indicators are sampled just before the first trade occurs after the α^H estimation interval. The regression also includes (in addition to a constant and two lags of the dependent variable) the control variables realised volatility (*rv*), trade size (*tsize*), trade duration (*dura*), unsigned trade imbalance (*imb*) and signed trade imbalance (*simb*). All variables are diurnally adjusted. See table 4 for computational details of the procedure. To obtain comparable estimates across stocks dependent variables and regressors are standardised by division by the sample standard deviations. The table reports OLS estimates averaged across stocks. *sig_pos* (*sig_neg*) counts the number of significant and positive (negative) coefficients. The significance level is 5%. Adjusted coefficients of determination are averaged across stocks. The regressions use on average 692 observations. Per activity quartile we have 953 (Q1), 744 (Q2), 516 (Q3), 537 (Q4) observations on average.

Table 6: Regression of trade informativeness (Hasbrouck-α) on liquidity measures and control variables.

|	book liquidity indicators					control variables							R^2_{adj}
	$spread$	$ap(50)$	$ap(200)$	$dpi50$	$dpi200$	rv	$tsize$	$dura$	imb	$simb$	$lag1$	$lag2$	
All stocks	0.043	-0.070	0.204	0.030	-0.098	0.023	-0.096	0.045	0.004	0.005	0.189	0.130	0.164
sig_pos (of 30)	11	0	22	2	0	3	0	7	1	0	27	26	
sig_neg	0	6	0	0	15	0	17	0	0	2	0	0	
Q1	0.064	-0.064	0.152	0.044	-0.081	0.048	-0.097	0.019	0.022	0.026	0.226	0.155	0.184
Q2	0.048	-0.088	0.208	0.044	-0.119	0.012	-0.115	0.047	0.006	-0.007	0.205	0.158	0.194
Q3	0.038	-0.040	0.193	-0.002	-0.074	0.004	-0.097	0.051	-0.009	0.000	0.161	0.097	0.124
Q4	0.020	-0.089	0.266	0.038	-0.118	0.033	-0.073	0.061	-0.003	0.005	0.167	0.111	0.156

Note: The trade informativeness measure α^H is regressed on the book liquidity measures *spread*, *ap50*, *ap200*, as well as *dpi50* and *dpi200* which denote the difference between the ask and the bid price concession for a hypothetical trade of €50,000 and €200,000, respectively. The liquidity indicators are sampled before the first trade occurs after the previous α^H estimation interval. The regression also includes (in addition to a constant and two lags of the dependent variable) the control variables realised volatility (*rv*), trade size (*tsize*), trade duration (*dura*), unsigned trade imbalance (*imb*) and signed trade imbalance (*simb*). These variables are computed using data from the previous α^H estimation interval. All variables are diurnally adjusted. See table 4 for computational details of the procedure. To obtain comparable estimates across stocks, the dependent variable and the regressors are standardised by division by the sample standard deviations. The table reports stock group averages of the OLS estimates. *sig_pos* (*sig_neg*) counts the number of significant and positive (negative) coefficients. The significance level is 5%. Adjusted coefficients of determination are averaged across stocks. The regressions use on average $T = 692$ observations. Per activity quartile we have 953 (Q1), 744 (Q2), 516 (Q3), 537 (Q4) observations on average.

Figure 1: Time-of-day patterns of the trade informativeness measure α_j^H.
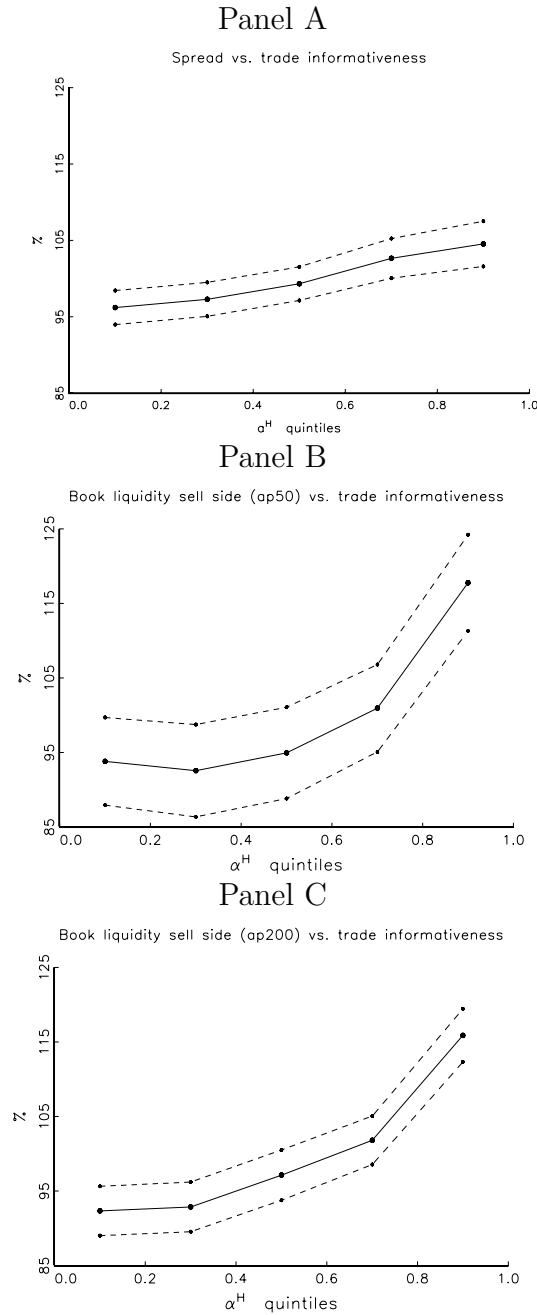
Left Panel: Q1 (most active) right panel: Q2
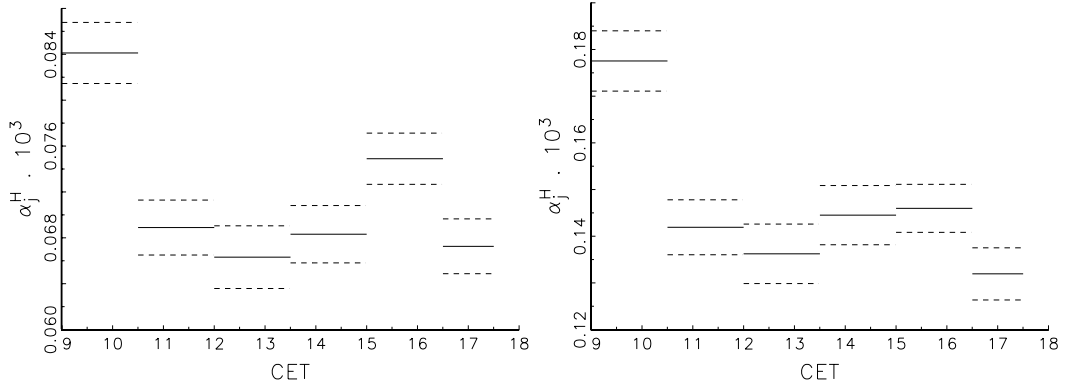
Left Panel: Q3 right panel: Q4 (least active)
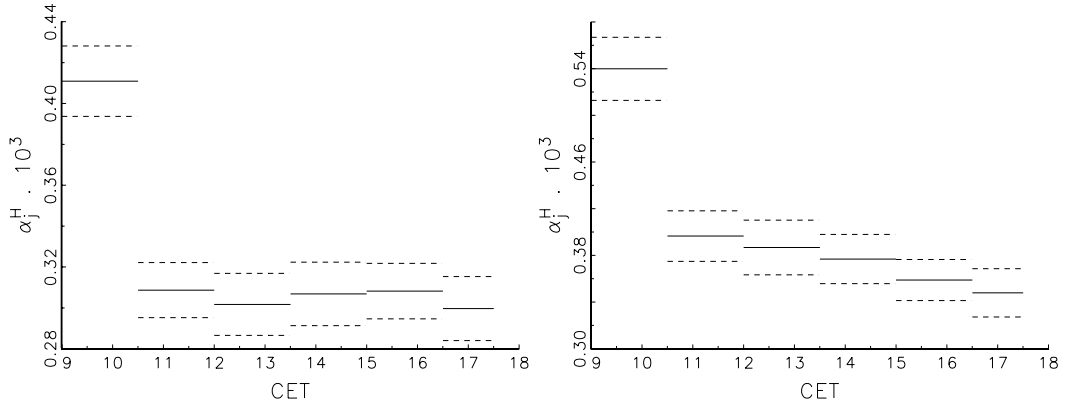
Note: The four panels of the figure show 90 minutes averages of the estimated trade informativeness measures α_j^H. The averages are computed over all α_j^H estimates in the respective 90 min. interval and over all stocks belonging to the respective trading activity quartile. The top left panel displays the results for the first quartile (most actively traded), the top right panel for the second quartile, the bottom left panel for the third quartile and the bottom right panel for the fourth quartile (least active). The dashed lines represent bounds of the 95% confidence interval.

Figure 2: Book liquidity measures versus trade informativeness.

Panel A

Spread vs. trade informativeness

Panel B

Book liquidity sell side (ap50) vs. trade informativeness

Panel C

Book liquidity sell side (ap200) vs. trade informativeness

Note: We pool time series of trade informativeness measures α^H (Hasbrouck-α) for the 30 stocks and sort them into quintiles. We then compute the quintile means of the book snapshot variables *spread*, $ap(50)$ and $ap(200)$. Quintile means are represented as dots connected with solid lines. The small dots connected with dashed lines are bounds of the 95% confidence intervals. Both α^H and book liquidity measures are diurnally adjusted (division by time-of-day mean) prior to the analysis. Diurnally adjusted book liquidity variables are multiplied by 100 to obtain percentages.

CFS Working Paper Series:

No.	Author(s)	Title
2011/08	**Joachim Grammig** **Erik Theissen** **Oliver Wünsche**	Time and the Price Impact of a Trade: A Structural Approach
2011/07	**Tullio Jappelli** **Mario Padula**	Investment in Financial Literacy and Saving Decisions
2011/06	**Carolina Achury** **Sylwia Hubar** **Christos Koulovatianos**	Saving Rates and Portfolio Choice with Subsistence Consumption
2011/05	**Gerlinde Fellner** **Erik Theissen**	Short Sale Constraints, Divergence of Opinion and Asset Values: Evidence from the Laboratory
2011/04	**André Betzer** **Jasmin Gider** **Daniel Metzger** **Erik Theissen**	Strategic Trading and Trade Reporting by Corporate Insiders
2011/03	**Joachim Grammig** **Eric Theissen**	Is BEST Really Better? Internalization of Orders in an Open Limit Order Book
2011/02	**Jördis Hengelbrock** **Eric Theissen** **Christian Westheide**	Market Response to Investor Sentiment
2011/01	**Mathias Hoffmann** **Michael U. Krause** **Thomas Laubach**	Long-run Growth Expectations and "Global Imbalances"
2010/26	**Ester Faia**	Credit Risk Transfers and the Macroeconomy
2010/25	**Ignazio Angeloni** **Ester Faia** **Roland Winkler**	Exit Strategies

Copies of working papers can be downloaded at http://www.ifk-cfs.de