Bid- and ask-side liquidity in the NYSE limit order book[☆]

Tolga Cenesizoglu

Alliance Manchester Business School and HEC Montréal

Gunnar Grass[*]

HEC Montréal

Abstract

We document significant differences between bid- and ask-side liquidity using 11 years of comprehensive NYSE limit order book data. First, the ask- but not bid-side liquidity of financial stocks deteriorates during the 2008 short selling ban. Second, bid-side (ask-side) liquidity decreases (increases) in lagged short- and long-term returns, indicating persistent contrarian behavior in limit orders. Third, liquidity commonality increases during the financial crisis, more so on the bid than on the ask side. Finally, ask- but not bid-side illiquidity predicts daily returns, while both forecast monthly returns. We discuss implications for researchers, investors, and policy makers.

Key words: Market Liquidity, Limit Order Book, Financial Crisis, Short Selling Ban, Asset Pricing
JEL classification: G12

[☆] The paper benefited substantially from helpful comments from Patrick Augustin, Christian Dorion, Jan Ericsson, Pascal Francois, Ruslan Goyenko, Alexandre Jeanneret, Ingrid Werner, and participants at the HEC-McGill Winter Finance Workshop. We thank Craig Holden and Stacey Jacobsen for providing their SAS code online. Financial support from the Fonds de Recherche du Quebec sur la Société et la Culture (FRQSC) is gratefully acknowledged.

[*] Corresponding author:
Gunnar Grass
Address: 3000, ch. de la Côte-Sainte-Catherine, H3T 2A7 Montréal, QC, Canada
Phone: (+1) 514 340 1540
Fax: (+1) 514 340 5632
Email: gunnar.grass@hec.ca

1. Introduction

Various studies show that liquidity provision in downward markets differs fundamentally from that in upward markets. Brunnermeier and Pedersen (2009) model mutually reinforcing market and funding liquidity spirals and show that these can explain sudden drops in liquidity as well as increases in commonality during market downturns. Hameed et al. (2010) empirically confirm the predicted links between market valuation and liquidity as well as its commonality and Chordia et al. (2001) report that aggregate market liquidity "plummets" in down markets. In contrast, the urgency of the demand for liquidity driving these results is less likely to occur when market values appreciate. As downward markets are characterized by excess selling of stocks, it is intuitive to expect that the described liquidity asymmetry between up- and downward markets translates into significant differences between buy- and sell-side liquidity.

However, most articles studying market liquidity do not make this distinction but employ liquidity measures which are based on the assumption that negative and positive order flows affect prices symmetrically.[1] An exception is the study of Brennan et al. (2012), who propose a way to estimate monthly market depth parameters for the buy and sell sides based on the studies of Glosten and Harris (1988) and Brennan and Subrahmanyam (1996). Brennan et al. (2013) decompose the Amihud (2002) measure into up- and downward liquidity to show that the latter is priced while the former is not. Almost all studies of market liquidity employ measures calculated from historical transaction prices or Level I bid and ask quotes.

In this paper, we disentangle buy- and sell-side liquidity by observing ask- and bid-side transaction costs for novel limit order book (LOB) data provided in the Thomson Reuters Tick History (TRTH) market depth files. TRTH provides millisecond timestamped snapshots of the LOB, including prices and quantities at its first ten levels on the ask and bid sides for each stock traded on the New York Stock Exchange (NYSE). Our data begin in 2002 and end in 2012. Given Amihud et al. (2012)'s broad definition of a liquid security as one

[1] For a summary of important liquidity measures, see Holden et al. (2014).

that can be traded in large amounts quickly and at a low cost, the LOB is a natural source of information concerning market liquidity. It allows one to observe – rather than estimate – the demand and supply schedules and thus trading costs of an investor seeking to buy or sell securities by placing a market order. To the best of our knowledge, we are the first to study comprehensive LOB data for the US equity market.[2]

To summarize the information embedded in the LOB of a stock, we propose a liquidity taker's marginal cost of immediacy (MCI) as a new measure of LOB liquidity and compute it for both, the bid- and ask-side. We define MCI_A (MCI_B) as the volume-weighted transaction costs (relative to the midprice) of an investor instantaneously accepting all ask (bid) offers in the first ten levels of the book, scaled by the total dollar volume he acquires (sells). We interpret MCI_A and MCI_B as measures of ask- and bid-side transaction costs and illiquidity. Additionally, we measure LOB imbalances MCI_{IMB} to examine asymmetries in order book liquidity. We do not introduce MCI to improve upon existing liquidity measures – in fact, MCI_A and MCI_B are highly correlated with the Amihud (2002) measure of liquidity, amongst others – but simply given the lack of a Level II liquidity measure that (i) is in line with the above definition of liquidity, (ii) combines the information on LOB prices and quantities, (iii) can be compared over time and across stocks and (iv) uses information embedded in all ten levels of the LOB.[3]

Our empirical analysis contributes to research on liquidity, market microstructure, and asset pricing in the following ways. First, we report details on the cleaning of the novel TRTH LOB dataset and show how it can be matched to other databases even without a common identifier. Second, we document and discuss

[2]Kalay and Wohl (2009) examine all orders submitted to the opening session of the Tel Aviv Stock Exchange. Cao et al. (2009) document a contribution of LOB data to price discovery for Australian data. Roesch and Kaserer (2014) study an aggregate measure of the LOB of German stocks - the Xetra Liquidity Measure. Marshall et al. (2011) rely on TRTH LOB data to study the liquidity of commodities. Fong et al. (2014) use TRTH data to analyze the performance of liquidity proxies computed for a global sample of equities but restrict their study to Level I data.

[3]The main intuition for our measure – capturing transaction costs per volume traded – corresponds to that of the Xetra Liquidity Measure (XLM) developed by the German Stock Exchange (see Roesch and Kaserer (2014)). However, the XLM measure fails to meet these criteria because it is computed for a specific trading volume, for example ten thousand or one million dollars. While the quote volume in the LOB of most NYSE stocks does not amount to one million dollars – and the XLM can accordingly not be computed for this amount – the LOB of the largest stocks exceeds one million dollars and the XLM measure does not capture all information in the book.

the relationship between Level II liquidity measures and the set of Level I measures computed from TAQ by Holden and Jacobsen (2014) to provide an understanding of the information content of the NYSE LOB. Third, we provide an overview of the evolution of market-wide LOB liquidity and imbalance. Average bid- and ask side-transaction costs MCI_B and MCI_A are 1.64 and 1.54 basis points per additional one thousand USD traded and were downward trending until the onset of the financial crisis, during which they spiked. The ask- but not bid-side liquidity of financial stocks deteriorates during the 2008 short selling ban, indicating heavy distortions in the order book caused by this regulatory intervention. Fourth, we examine the factors driving bid- and ask-side liquidity. The economically most important determinants of firm-level order book liquidity are turnover, firm size, volatility, and competition for market making, which affect bid- and ask-side liquidity in similar ways. However, MCI_A (MCI_B) decreases (increases) in lagged short- and long-term returns, indicating persistent contrarian behavior in limit orders. Turnover, volatility and returns also affect market-wide average MCI_A and MCI_B, which may explain the common component in liquidity. Liquidity commonality increases during the financial crisis, more on the bid than on the ask side. Fifth, we provide evidence indicating that order book liquidity predicts future returns through two channels, short term price pressure from imbalances and liquidity premia. MCI_A and MCI_{IMB}, but not MCI_B, predict daily returns, while both MCI_A and MCI_B are positively related to future monthly returns.

The remainder of the paper is organized as follows. Section 2 introduces the marginal cost of immediacy as a new LOB liquidity measure. Section 3 describes our sample selection, compares Level II to Level I liquidity measures and displays time series of LOB liquidity and imbalance. Section 4 identifies the determinants of LOB ask versus bid liquidity at the firm and aggregate levels and documents the existence of and changes in liquidity commonality. Section 5 then documents that our MCI measures predict daily and monthly stock returns. Finally, Section 6 concludes.

2. Measuring LOB Liquidity

Our main objective in this paper is to disentangle bid and ask side liquidity. To do so, we analyze comprehensive LOB data for the US equity market. Given the limited number of studies using any type of LOB data, the lack of a well established measure of Level II liquidity is not surprising. In this section, we develop an LOB liquidity measure in the spirit of Grossman and Miller (1988), who motivate their model of the supply and demand of immediacy by noting that "the cost of trading immediately rather than delaying the order, particularly when the order is a large one [...] is the essence of market liquidity". More recently, Chacko et al. (2008) model transaction costs as the price of immediacy, which is zero if markets are perfectly liquid and otherwise increases in the traded quantity. We define our liquidity measure accordingly as the value-weighted transaction costs of an investor who demands immediacy by instantaneously accepting all ask (bid) offers in the first ten levels of the order book, scaled by the total dollar volume he acquires (sells). We label our measure the *"Marginal Cost of Immediacy"* and use it to separate bid- and ask-side liquidity using high-frequency TRTH limit order book data as detailed in the following.

The TRTH Market Depth (MD) NYSE file used in this study contains millisecond timestamped snapshots of the LOB including ask and bid prices and quantities for the first ten levels of the book. Figure 1 shows a sample snapshot of the LOB for IBM on July 2nd, 2007 at 3 PM. The left and right sides show bid (B1 to B10) and ask (A1 to A10) information for the first ten levels of the book, respectively. The upper graph separately plots the order price (line) and order size (bars). In the lower panel, both dimensions are combined.

[Figure 1 about here.]

The calculation of MCI_A and MCI_B can be illustrated using the transformed order book shown in the lower graph of Figure 1, in which the logarithm of the volume-weighted average price of the ask (bid) side scaled by the midprice, $VWAPM_{A,L}$ ($VWAPM_{B,L}$), is plotted against the cumulative dollar volume at level L of the ask (bid) side, $Vlm_{A,L}$ ($Vlm_{B,L}$). More precisely, $VWAPM_{A,L}$ is defined for a given snapshot of the LOB as follows:

$$VWAPM_{A,L} = \ln\left(\frac{VWAP_{A,L}}{M}\right), \tag{1}$$

where

$$VWAP_{A,L} = \frac{Vlm_{A,L}}{\sum_{l=1}^{L} Q_{A,l}} \tag{2}$$

and $Vlm_{A,L} = \sum_{l=1}^{L} P_{A,l} \times Q_{A,l}$. The price and quantity available at the l^{th} level of the ask side are denoted by $P_{A,l}$ and $Q_{A,l}$, respectively, and M is the midquote price defined simply as the average of $P_{A,1}$ and $P_{A,2}$. For the bid side, $VWAPM_{B,L}$ is defined analogously.

$VWAPM_{A,L}$ ($VWAPM_{B,L}$) thus corresponds to the transaction cost of buying (selling) all stocks available up to level L in the book, expressed as a log return relative to the midquote price. Scaling this by the total dollar volume he acquires (sells) yields our measure of marginal transaction costs, the marginal cost of immediacy, MCI_A (MCI_B), for the ask (bid) side. MCI_A and MCI_B correspond to the inverse of the slope of the ask and bid sides in this transformed LOB and can be computed as:

$$MCI_A = \frac{VWAPM_{A,10}}{Vlm_{A,10}} \tag{3}$$

and

$$MCI_B = \frac{-VWAPM_{B,10}}{Vlm_{B,10}}, \tag{4}$$

for the ask and bid sides, respectively. We measure MCI_A and MCI_B in basis points (bp) per 1,000 US dollars. The measures are straightforward to interpret. For instance, $MCI_A = 0.2$ means that an immediacy-demanding investor who purchases an additional USD 1,000 by accepting sell orders shown in the ask schedule of the book will face additional transaction costs of 0.2 bp. If he purchased an additional USD 1,000,000, transaction costs would increase to 2.0%.

In addition to the level of MCI, we study its bid-ask imbalance, defined as the difference between MCI_A and MCI_B, scaled by their sum:

$$MCI_{IMB} = \frac{MCI_A - MCI_B}{MCI_A + MCI_B}. \tag{5}$$

As both MCI_A and MCI_B are positive by construction, MCI_{IMB} lies between -1 and +1. An MCI_{IMB} of zero implies that accepting all sell orders is as expensive as accepting all buy orders. Values below zero mean that the marginal cost of accepting ask orders (buying) is lower than the marginal cost of accepting bid orders (buying). The opposite is true for positive values. For MCI_{IMB} equal to +1 (-1), all of a liquidity taker's transaction costs are due to buying (selling).

MCI_A and MCI_B are similar to some liquidity proxies computed from Level I data. They are inversely related to the slope of the ask and bid sides of the graph shown in the lower panel of Figure 1. The steeper the slopes are, the lower MCI_A and MCI_B are. Closely related, Kyle (1985)'s λ can be estimated as the slope coefficient from a regression of price change on signed volume. Extending the work of Glosten and Harris (1988), Brennan et al. (2012) estimate monthly lambdas separately for the buy and sell sides. In contrast to these liquidity proxies, MCI can be observed and does not need to be estimated from time series of Level I data. MCI is therefore not affected by potential estimation errors and can be computed at each snapshot of the LOB, which makes it especially useful for research on bid- and ask-side liquidity at a daily or higher frequency.

To the best of our knowledge, our MCI measures are the first measures of LOB liquidity and imbalance meeting the following criteria. First, MCI is in line with the broad definition of liquidity offered by Amihud et al. (2012), according to which a liquid security is one that can be bought and sold in large amounts at a low cost.[4] Second, it combines the information on LOB prices and quantities. Third, it exhausts information from all levels of the LOB. Finally, MCI can be compared over time and across stocks, enabling our analysis of comprehensive panel data. As discussed in Appendix A-3, while MCI is similar to the "Xetra Liquidity Measure" published by the German Stock Exchange to describe the LOB liquidity, the latter is not suitable for studying a heterogenous cross-section of stocks, as is done in this article.

We include the following additional measures based on the TRTH dataset in our overview of the dataset presented subsequently. We compute the *Percent Quoted Spread* ($PrctQuotedSpread_L$) as the bid-ask spread scaled by the midquote price. The *Aggregate Dollar Volume* (Vlm_L) of the book is the sum of the cumulative volume at the same level L of the bid and ask sides. *Volume Imbalances* are the difference between ask and bid volume scaled by their sum. We compute these measures for levels $L = 1$ and $L = 10$, and they serve two purposes. First and foremost, the correlation between those based on $L = 1$ using TRTH and the same variable computed using TAQ provide information on the validity of the TRTH dataset. More precisely, a high correlation would suggest that Level I information from TRTH is similar to that from TAQ. Second, those based on $L = 10$ provide some information on the cost and quantity dimensions separately.

3. Dataset and Summary Statistics

3.1. Data Sources and Sample Selection

This study aggregates high-frequency stock market data from the Thomson Reuters Tick History (TRTH) and the Trade and Quote (TAQ) databases into a joint dataset of daily observations that is complemented by data from the Center for Research in Security Prices (CRSP) and Compustat. The primary sources of data used in this study are the TRTH Market Depth (MD) New York Stock Exchange (NYSE) files, which

[4] Widespread measures consistent with this general definition include those proposed by Roll (1984), Ho and Macris (1984), Choi et al. (1988), Glosten and Harris (1988), Huang and Stoll (1996), Lesmond et al. (1999), Amihud (2002), Pastor and Stambaugh (2003), Hasbrouck (2004), Holden (2009), and Feldhuetter (2011).

include comprehensive high-frequency LOB data. They comprise the bid and ask prices and aggregate order volumes for ten levels for each side of the book (bid and ask). The snapshot of the LOB for a single stock at a specific point in time thus consists of 40 data points. Each snapshot is identified by a Reuters Instrument Code (RIC) and a millisecond timestamp. As soon as there is a change in price or quantity at any level of the book due to a newly placed, withdrawn, or executed order, a new snapshot of the entire book is created. In other words, we can observe the LOB for all stocks in our sample with millisecond precision.[5] Our sample period begins in January 2002 – when the NYSE began making level-two LOB data available to market participants outside the trading floor and TRTH MD data begin – and ends in December 2012.[6]

The initial dataset includes 52.44 billion snapshots. As detailed in Appendix A-2, our sample selection criteria are similar to these used by Korajczyk and Sadka (2008). We require that each snapshot include price and volume data for all levels and that prices increase monotonically throughout the book. We further delete observations with Level 1 (Level 10) bid-ask spreads above 25% (250%), midprices below $1 or above $1000. Next, we eliminate securities for which we are unable to establish a link to data reported in the Center for Research in Security Prices (CRSP) files, using the matching procedure described in Appendix A-1, and all securities other than ordinary common shares. Finally, deleting all firm-days with fewer than 100 snapshots in the TRTH data or fewer than 100 trades in the matched TAQ dataset as well as observations with incomplete CRSP or Compustat data results in a sample of 31.19 billion observations reported for 2,103 stocks over 2,740 trading days and 3.37 million stock-day observations. Chordia et al. (2002) study the relationship among liquidity, order imbalances, and returns and argue that daily time intervals balance the trade-off between reducing problems related to very-high-frequency data and capturing short-term effects

[5]Although the TRTH MD data are also available for stocks listed on NASDAQ, we follow Korajczyk and Sadka (2008) and restrict our analysis to stocks listed on the NYSE due to differences in trading mechanisms between the two exchanges.

[6]We exclude days for which the number of book snapshots per trading hour is less than 50% of the monthly average, namely January 28th, February 1st, October 31st and November 20th, 2002, January 21st and October 15th, 2003, January 13th, 2004, and May 4th, 2009. TRTH confirms technical problems related to data collection for these days.

among the variables of interest. In line with their argument, we aggregate all TRTH and TAQ liquidity measures on a daily basis by computing their equally weighted average.[7]

Of course, our dataset only reflects a part of all the liquidity provided by market participants. First, the TRTH data do not include hidden orders. Second, the TRTH comprises only NYSE limit orders and not limit orders from other markets, unlike TAQ data. Our liquidity measures might therefore overstate the actual transaction costs. However, as argued in Fong et al. (2014) and shown below, TRTH passes several integrity checks and provides useful information on the true cost of trading.

3.2. Level I Measures of Liquidity

In this section, we briefly discuss some Level I measures computed using TAQ data, which mainly serve as a validity tool for the TRTH data, as mentioned at the end of Section 2, and leave the details concerning the computation of these measures to Appendix A-4.

We calculate all of the measures discussed in Holden and Jacobsen (2014). Given the very high correlations between some of their measures, we report results only for the percent quoted spread in basis points (*PrctQuotedSpread*), percent effective spread in basis points (*EffectiveSpreadPercent*), total dollar depth available at both the best bid and ask prices in millions of dollars ($TotalDepthDollar$), the total dollar depth imbalance between bid and ask sides ($TotalDepthDollarImbalance$), the percent price impact in basis points ($PercentPriceImpact$), and the buy-to-sell order imbalance ($OrderImbalance$). Whenever necessary, we use the Lee and Ready (1991) algorithm to classify orders into buys and sells. All measures are computed from the monthly TAQ files using the "Interpolated Time Technique" proposed by Holden and Jacobsen (2014). In addition to their measures, we also estimate buy- and sell-side illiquidity following Brennan et al. (2012). Their buy and sell lambdas measure the market impact of trading on the ask and bid

[7] We also considered a time-weighted average, i.e., assigning greater weight to snapshots that were in effect for a longer period of time, and our results remain similar.

side, respectively, and are thus conceptually similar to MCI_B and MCI_A. Finally, we also consider the common Amihud (2002) illiquidity measure (*AmihudIlliq*) and the annualized realized volatility (*RealizedVol*).

3.3. Descriptive Statistics and Correlations

Examining the relationship between Level I and Level II measures computed from information beyond the best bid and ask provides initial insights into the information content of Level II data. Table 1 presents descriptive statistics for variables computed from TRTH Level II LOB data (Panel (A)) and TAQ Level I data (Panel (B)).

[Table 1 about here.]

We compute TRTH (TAQ) measures for 3.37 million firm-days from 9,250 LOB snapshots (6,666 trades) per firm-day on average. The average (median) percent quoted spread equals 20.3bp (12.0bp) for TRTH data and 16.1bp (10.5bp) for TAQ data. The difference between these can be explained by the fact that the TAQ bid-ask spread is computed for the National Best Bid and Offer (NBBO) quotes, where the national best bid (offer) is the highest (lowest) quote available across all US stock exchanges, whereas the TRTH bid-ask spread is computed based on NYSE data alone. The volume available at the Level-1 bid and ask equals 53.6 million USD on average, representing only a small fraction of the volume in the first ten levels of the book, which equals 913.9 million USD. However, trading larger volumes instantaneously is costly: The average Level-10 bid-ask spread equals 2.8%. The dollar depth at the NBBO is substantially lower than that observed for NYSE orders. For both TRTH and TAQ data, we observe that the ask-side volume is slightly larger than the bid-side volume, although this difference is not statistically significant. Marginal transaction costs *MCI* equal 3.2 bp per 1,000 USD on average for a roundtrip trade. In line with the higher depth of the ask side, selling is more expensive than buying. The average transaction cost imbalance equals 47.3%. Overall, the level of and variation in TAQ measures are comparable to those reported in Holden and Jacobsen (2014). All measures except the LOB imbalance measures exhibit substantial variation and skewness.

[Table 2 about here.]

We now turn our attention to the correlations presented in Table 2. TAQ *PrctQuotedSpread* and TRTH $PrctQuotedSpread_{L=1}$ are 90% correlated. Similarly, $TotalDepthDollar$ based on TAQ and its analog based on TRTH, $Vlm_{L=1}$, are also highly correlated at 87%. These results suggest that Level I information from TRTH is similar to that from TAQ and confirms the integrity of the TRTH dataset reported in Fong et al. (2014). The remaining results are in line with intuition. Higher depth is accompanied by lower spreads. Marginal transaction costs MCI are positively related to spreads and negatively related to volume. The correlation between MCI_A or MCI_B and the widespread Amihud (2002) measure of market impact amounts to 74%. Imbalances in depth and MCI exhibit a substantial negative relationship as expected. Correlations between imbalance measures and illiquidity measures are low, indicating that order imbalances capture different information.

3.4. Time Series of Average Bid- and Ask-side Liquidity

Before analyzing the determinants of liquidity in the next section, we present evidence on the variation of market-wide LOB liquidity over time. We first examine the evolution of transaction costs and their imbalances over the entire sample period and then focus on changes in liquidity around the 2008 short selling ban. Figure 2 displays monthly averages of ask- and bid-side transaction costs MCI_A and MCI_B. The upper plot shows value-weighted and the lower plot equally weighted averages. We observe improvements in market liquidity (decreases in transaction costs) prior to the financial crisis. In part, this decrease can be attributed to the end of the market downturn induced by the burst of the tech bubble in 2000 and lasting until early 2003. In relative terms, the decrease in transaction costs was larger for firms with a high market capitalization, as reflected by the steeper decrease in the value-weighted relative to the equally weighted average. We attribute the improvements in trading costs prior to the financial crisis to the long-run trend towards increased market liquidity documented by Brennan et al. (2012), amongst others. Trading costs jump sharply in the aftermath of the bankruptcy of Lehman Brothers and remain at high levels until market valuations reached the bottom in early 2009. A similar pattern can be observed following the burst of the dot-com bubble. Liquidity began improving following the market turnaround in January 2003. Overall, transaction costs are high when market valuations are low and volatility is high.

[Figure 2 about here.]

The time series of average imbalances in the LOB shown in Figure 3 are less straightforward to interpret. Consistent with the finding that sell lambdas tend to be higher than buy lamdbas reported by Brennan et al. (2012), buying in the LOB (accepting ask orders) is less expensive than selling for the majority of the sample period. The gap between MIC_A and MCI_B is highest at the beginning of our sample period and decreases until the onset of the financial crisis, when, during a very brief period in the fall of 2008, average MCI_{IMB} values reach peak levels above 0.05. As detailed in the following, this increase is driven by substantial decreases in the ask-side liquidity of financial stocks due to the short selling ban imposed from September 19th to October 8th, 2008 following the collapse of Lehman Brothers.

[Figure 3 about here.]

Figure 4 displays daily averages of the transaction cost imbalance MCI_{IMB} around the ban. The upper figure shows value-weighted and the lower figure equally weighted averages. The effect of the short selling ban on the limit order imbalance MCI_{IMB} of financial stocks is dramatic. On the first day of the ban, ask- and bid-side costs are approximately equal. Ask-side costs increase sharply following the implementation of the short selling ban. While the ban is active, the volume-weighted average ask-side trading cost is approximately four times as high as the bid-side cost. Immediately after the end of the ban, MCI_{IMB} declines to approximately zero. The effect is stronger for large than for small firms. To the best of our knowledge, we are the first to document this finding. It adds to the discussion on the usefulness of the short selling ban, which has been questioned by academics and practitioners alike. [8]

[Figure 4 about here.]

To further decipher the sources of the effect, Table 3 documents changes in Level-10 spreads and volumes around the 2008 short selling ban for financial services firms (Panel A) and all other firms (Panel B). For financial firms, the aggregate bid- and ask-side volume observable in the first ten levels of the book does not change significantly. Only the level-10 mid-to-ask spread increased significantly, driving the reported results. In addition to the substantial decline in the ask-side liquidity of financial stocks, we also find significant decreases in ask-side liquidity for non-financial firms. This extends the evidence on spillover effects provided by Boehmer et al. (2015).

[Table 3 about here.]

4. Determinants of Bid versus Ask Liquidity

A tremendous amount of empirical and theoretical research has been dedicated to identifying the determinants of symmetric measures of illiquidity such as the bid-ask spread. However, to the best of our knowledge, evidence on differences in the determinants of buy- and sell-side liquidity is limited to Brennan et al. (2012). We complement their research by first comparing firm-level determinants of bid and ask LOB liquidity using panel data. We next analyze the time-series determinants of market-wide (average) bid versus ask liquidity. Finally, we study commonality in bid and ask liquidity to understand whether firm-level liquidity is partly determined by market liquidity.

[8] See Beber and Pagano (2013), as well as Christopher Cox's, telephone interview with Reuters, December 31, 2008.

4.1. Determinants of Firm-Level Liquidity

To identify and compare the firm-level determinants of bid- and ask-side LOB liquidity, we jointly estimate seemingly unrelated regressions (SUR) of logged MCI_A and MCI_B measures and test for the difference in coefficients using a Wald test. We follow Aragon and Strahan (2012) and use logged liquidity as a dependent variable in the first two regressions because both MCI measures are heavily skewed. Furthermore, we present results of regressions with MCI_{IMB} as the dependent variable. Explanatory variables include traditional determinants of the bid-ask spread, proxies for trade direction, and a set of control variables. In all three regressions, standard errors are clustered at the firm level following Petersen (2009).

Table 4 shows our results. In addition to the regression estimates, we report indicators of the economic significance of the relationship between the dependent and explanatory variables in the last three columns. They display the product of coefficient estimates and the standard deviation for each explanatory variable. As the first two regression models are log linear (the dependent variables are $logMCI_A$ and $logMCI_B$), this product corresponds to the log change in MCI_A and MCI_B given a one-standard-deviation increase in the explanatory variable. For the third model, it simply indicates how much MCI_{IMB} changes in absolute terms given this increase.

[Table 4 about here.]

Our MCI measures capture the cost of trading in the LOB, which consists of a collection of bid-ask quotes. It is thus natural to use variables in our regressions that have traditionally been employed to explain variations in the bid-ask spread. These include a market maker's order processing costs, his inventory holding costs (Garman (1976), Stoll (1978), Amihud and Mendelson (1986), Ho and Stoll (1981)) and his costs of adverse selection due to the risk of being picked off by an informed trader (Glosten and Milgrom (1985), Kyle (1985)). In addition, several studies, including Chacko et al. (2008), argue that market makers can use their power to extract rents from selling immediacy if market making is not perfectly competitive. We proxy for these four determinants as follows.[9] We measure order processing costs as $Turnover$, defined as the log of the daily dollar trading volume reported in CRSP. Not surprisingly, the relationship between $Turnover$ and our illiquidity measures $logMCI_A$ and $logMCI_B$ is negative and highly significant. Furthermore, $Turnover$ has by far the highest economic impact on trading costs. A one-standard-deviation decrease in $Turnover$ results in an increase (stated as log change) of 97.9% (101.9%) in MCI_A (MCI_B). As indicated by differences in the coefficients, the impact of $Turnover$ is higher on bid-side liquidity than on ask-side liquidity. This is also reflected by a positive and significant relationship between $Turnover$ and MCI_{IMB}. Our proxy for inventory holding costs is realized volatility $RealizedVol$, which we compute from intraday TAQ data. In line with expectations, the relationship between both of our illquidity measures and $RealizedVol$ is positive, statistically significant and economically meaningful. A one-standard-deviation increase in volatility increases MCI_A and MCI_B by approximately 42%. $RealizedVol$ does not appear to drive difference in bid- versus ask-side liquidity; the difference between coefficients is insignificant. The relationship between realized volatility and MCI_{IMB} is negative but only significant at the 10% level. Given the lack of more detailed information, we capture $Competition$ between market makers as one minus the Herfindahl index measuring the dispersion of trading volume across exchanges based on TAQ data. Indeed, we find a statistically and economically significant negative relationship between $Competition$ and trading costs. We further observe that its impact on bid-side liquidity is substantially higher than that on ask-side liquidity together with a positive and statistically significant coefficient in the MCI_{IMB} regression. A potential explanation is that market makers are able to extract higher rents from the demand for immediacy on the sell (bid) side because the need to sell shares quickly is arguably more widespread than the urge to buy them. Increased competition could accordingly reduce the liquidity provider's rent more on the bid than on

[9]For an overview of empirical analyses of the bid-ask spread using these and alternative proxies, see Bollen et al. (2004).

the ask side. Finally, we include the Easley et al. (1996) probability of informed trading (PIN) estimated based on the approach described in Lin and Ke (2011) and absolute order imbalance (*AbsOrderImbalance*) as proxies for the cost of adverse selection. The calculation of PIN is detailed in Appendix A-5.[10] Again in line with expectations, we find that MCI increases in adverse selection costs. While the relationship is statistically significant, the factor's economic importance is relatively low.

Our next set of variables labelled "Trade Direction" captures whether a stock is or has been under directional trading pressure. It includes contemporaneous order imbalance (*OrderImbalance*) and return (*Ret*), the previous day's return (*RetLag*), as well as the Campbell et al. (2008) long-run excess return until the day before (*ExretAvg*). While the economic impact of these measures on bid- and ask-side liquidity is lower than that of the traditional determinants, all of them exhibit a significant and negative relationship with $logMCI_A$ and a significant and positive relationship with $logMCI_B$ and are, in economic terms, the most important determinants of imbalance in LOB liquidity, MCI_{IMB}. The higher lagged or contemporaneous returns and the higher the buy-sell order imbalance, the higher ask-side liquidity and the lower bid-side liquidity. Loosely speaking, investors provide more bid and less ask side liquidity during or following falling stock prices, while the opposite holds when stock prices rise. This evidence for contrarian behavior by liquidity providers complements studies documenting contrarian behavior by liquidity takers. For instance, Chordia et al. (2002) document increases (decreases) in the market buy-minus-sell order imbalance following market declines (increases). However, while they report that overall liquidity reduces in order imbalance independent of its direction, we are able to separate the effects on bid- and ask-side liquidity. Our findings are in line with the LOB model of Rosu (2009), which predicts that a sell order will decreases bid prices more than ask prices.

[10]We compute order imbalance as the difference between the number of Lee and Ready (1991) buy minus sell orders scaled by their sum. We include the daily available *AbsOrderImbalance* in addition to PIN, as the latter needs to be estimated from longer time series – we use data from the preceding month.

Finally, our set of control variables includes logged market capitalization ($Size$), the Campbell et al. (2008) measures log stock price winsorized at \$15, market to book ratio (MB), leverage ($Lever$), profitability ($Profit$), and cash holdings ($Cash$), a dummy for financial services firms (FS), one for days during which the 2008 short selling ban was active (*Ban*), and their interaction. Liquidity increases in $Size$, which is the second most important determinant of $logMCI_A$ and $logMCI_B$ in economic terms. The coefficients of most other controls are not surprising. It is noteworthy, however, that the illiquidity of non-financial services firms decreased during the 2008 short selling ban, indicated by the *Ban* coefficient. For financial stocks, this increase was stronger on the ask than on the bid side. These findings suggest that effects from the short selling ban on financial stocks spilled over to unregulated non-financial stocks. They are in line with the evidence for illiquidity spillovers in regulatory experiments documented in Boehmer et al. (2015).

4.2. Time-Series Determinants of Aggregate Liquidity

Chordia et al. (2000) examine liquidity commonality and point to the need for research on the macro determinants of liquidity. Before verifying the existence and significance of commonality in LOB liquidity in the next section, we aim to provide an understanding of the factors driving market-wide bid and ask liquidity and their imbalance in this section. As documented in Table 5, we regress the daily time series of equally and value-weighted averages of our MCI measures MCI_A, MCI_B and MCI_{IMB} on independent variables comprising those used by Brennan et al. (2012) to analyze the time series of their monthly buy- and sell-side measures of liquidity. Their variables include the VIX implied volatility index (VIX), the ratio of the number of stocks with a positive contemporaneous return to that with a negative return (*UpDown Ratio*), today's and yesterday's market return (Ret_M and $Ret_M Lag$), as well as the TED spread measuring funding liquidity (*TED*). We add the log of the market-wide dollar trading volume measured in billions of USD ($Turnover_M$), as turnover is the most important determinant of LOB liquidity at the firm level, and a dummy variable for the 2008 short selling ban on financial stocks (*Ban*). Market returns and volume are computed from stocks listed on the NYSE, AMEX and NASDAQ. As in the previous section, we assign these variables to the categories “Traditional Determinants”, “Trade Direction”, and “Control Variables”. As opposed to firm-level MCI measures, their averages are not plagued by heavy skewness, and we therefore do not estimate log-linear regressions but use the non-transformed MCI measures as dependent variables. As in the previous section, we report indicators of the economic significance of the relationship between dependent and explanatory variables, which are displayed in Panel B of Table 5. It reports the product of coefficient estimates and the standard deviation for each explanatory variable. In contrast to the previous section, none of our dependent variable is logged. These statistics thus indicate how much the dependent variable changes in absolute terms given a one-standard-deviation increase in the explanatory variable.

[Table 5 about here.]

Our findings indicate that the macro equivalents to our most important determinants of firm-level MCI measures are indeed important determinants of aggregate LOB liquidity and imbalance. Again, turnover ($Turnover_M$) and volatility (VIX) drive aggregate LOB liquidity. Their relationship with equally and value-weighted MCI_A and MCI_B is statistically highly significant. In addition, they are the two most economically important determinants of LOB liquidity. The impact of $Turnover_M$ on both equally and value-weighted bid side liquidity MCI_B is substantially more negative than on MCI_A. Results for the relationship between measures of trade direction, including contemporaneous returns, lagged returns, and the *UpDown Ratio*, are mixed. Their relationship with MCI_A and MCI_B varies across regression specifications, and even statistically significant relationships are of low economic importance. However, consistent with our firm-level results, their impact on MCI_B is higher than that on MCI_A and all of them are negatively related to equally and value-weighted LOB imbalance MCI_{IMB}. When markets rise, bid-side liquidity deteriorates more (or improves less) than ask-side liquidity. This indicates that the previously documented contrarian behavior of liquidity providers at the firm-level carries over to the aggregate market.

Surprisingly, while (not reported) bivariate correlations between the TED spread and all MCI measures are positive, the relationship becomes negative after controlling for other factors in three out of four regression specifications. For our dataset and at an aggregate level, we can therefore not confirm the hypothesis that limited funding liquidity during the financial crisis drove transaction costs, as would be expected from Brunnermeier and Pedersen (2009), for instance. The differential impact of the 2008 short selling ban on bid versus ask liquidity displayed in Figure 4 remains significant in this multivariate setting, as indicated by the significant difference between the *Ban* coefficients and the positive and significant relationship between *Ban* and MCI_{IMB}.

Overall, our results on the relationship between trading costs and explanatory variables are in line with our expectations based on the predictions of microstructure models and the study of aggregate market liquidity by Chordia et al. (2001). However, neither the theoretical nor the empirical literature provides much guidance on whether the explanatory variables should have any differential relationship with the ask and bid sides. The only exception is Brennan et al. (2012), who analyze differences in the determinants of monthly Level I estimates of buy- and sell-side liquidity. Our results on the time-series determinants of aggregate liquidity complement theirs in that they are based on daily Level II data comprising the financial crisis, confirm the differential impact of the 2008 short selling ban on bid versus ask liquidity, and highlight the substantial economic impact of $Turnover_M$ and VIX on liquidity. As volatility, turnover, and returns drive LOB liquidity both at the firm and aggregate levels, a natural next step is to examine the prevalence of commonality in liquidity.

4.3. Commonalities in Bid and Ask Side Liquidity

In this section, we study commonalities in liquidity variables. As Korajczyk and Sadka (2008) note, the absolute level of illiquidity only requires a small premium to compensate investors for higher transaction costs. In contrast, systematic comovements of individual with market-wide liquidity can justify substantial liquidity premia. Accordingly, commonality in market liquidity has important implications for asset pricing. We contribute to studies confirming the prevalence of commonality by documenting differences in bid-versus ask-side commonality and by studying it's time variation around the financial crisis.[11]

4.3.1. *Estimation*

Commonality in LOB liquidity implies covariation between stock-specific and aggregate MCI measures. As in Chordia et al. (2000), we estimate the following market model of liquidity for each stock:

$$\Delta LIQ_{i,t} = \alpha_i + \beta_{1,i}\Delta LIQ_{M,t} + \beta_{2,i}\Delta LIQ_{M,t-1} + \beta_{3,i}\Delta LIQ_{M,t+1} + \beta_{X,i}\mathbf{X}_t + \varepsilon_{i,t}, \tag{6}$$

where $\Delta LIQ_{i,t}$ is the percentage change in stock i's liquidity LIQ from trading day t-1 to t and $\Delta LIQ_{M,t}$ the contemporaneous change in the cross-sectional average of LIQ. $\mathbf{X}_t$ represents control variables including lagged, contemporaneous and lead market returns, as well as the relative change in the realized volatility of stock i. Studying changes rather than levels reduces econometric problems due to the persistence of liquidity. When computing average market liquidity LIQ_M, we exclude stock i. This avoids an overestimation of commonality due to noisy data and outliers. Moreover, this approach implies that the average beta of all stocks does not necessarily equal one. As highlighted by Chordia et al. (2000), although this adjustment has a minor effect on the explanatory variables and thus the coefficient for each stock, it can substantially affect average betas.

4.3.2. *Commonality in the Pooled Sample*

Table 6 presents the results of stock-level regressions using data from the entire sample period. Commonality is documented for bid- and ask-side LOB liquidity and its imbalance. In line with Chordia et al.

[11] Studies documenting and explaining different features of commonality include Chordia et al. (2000), Coughenour and Saad (2004), Korajczyk and Sadka (2008), Kamara et al. (2008), Comerton-Forde et al. (2010), and Karolyi et al. (2012).

(2000), the results confirm the existence of commonality in both MCI_A and MCI_B. The sum of concurrent, lead and lag coefficients is positive and significant, and market liquidity contributes 1.6% (1.5%) to the explanatory power of MCI_B (MCI_A) regressions on average. While this difference in marginal explanatory power ΔR^2 is insignificant, the economic magnitude of commonality is significantly larger on the bid than on the ask side. The sum of the three market illiquidity coefficients equals 0.62 for the former and 0.54 for the latter. An intuitive explanation for this is that market participants' need to sell is likely to co-move more systematically than their need to buy, especially in times of economic distress. Market sells (which are matched with and thus eliminate limit buy orders) are then likely to co-vary more systematically than market buys. Similarly, the number of placed limit buy (bid) orders is then likely to decrease in a more systematic fashion than is true for limit sells. To further investigate this hypothesis, the next section examines changes in the level of commonality around the financial crisis.

Liquidity imbalances MCI_{IMB} show no signs of commonality. On average, none of the three coefficients is significant. The median ΔR^2 is zero. This indicates that the strong positive relationship between MCI_{IMB} and stock returns documented in Section 5 is not due to the pricing of a systematic liquidity factor but rather reflects temporary and idiosyncratic price pressure.

[Table 6 about here.]

4.3.3. Time Series of Commonality

In this section, we explore variations in commonality over time. Given the lack of commonality in MCI_{IMB}, we do not report time series results for this variable but instead focus on MCI_A and MCI_B. Table 7 presents the results of the stock-level regressions described above estimated using data from each semester instead of the full sample period. Each row in Table 7 shows averages and statistics from a t-test for the sum of all three market liquidity coefficients (Sum), together with average ΔR^2. During the

2008 financial crisis, LOB liquidity of firms in the financial services industry was affected by factors less relevant for industrial firms, including the short selling ban on financial stocks discussed in Section 3. We therefore report commonality results for the subsample of financial services firms (SIC codes 6000-6999) and industrial firms (all others) separately.

[Table 7 about here.]

We find substantial time variation in liquidity commonality. Commonality increases for both industrial and financial services firms on the bid and ask side during the financial crisis, peaking in the second half of 2008. In line with our hypothesis stated in the previous section, ΔR^2s are significantly higher on the bid than on the ask side. Our results indicate that bid-side liquidity is more prone to being driven by systematic factors in times of crisis than is ask-side liquidity. This finding offers one explanation for why the premium for sell-side liquidity is higher than that for buy-side liquidity, as reported by Brennan et al. (2012) and Brennan et al. (2013). In contrast to the peak of commonality at the core of the financial crisis at the end of 2008, the substantial increase in commonality in the second half of 2009 is puzzling, and we leave it to future research to identify its causes.

5. Predictive Power of Bid versus Ask Liquidity

In this section, we analyze the power of our MCI measures to predict stock returns. We first make predictions concerning the channels through which we expect LOB liquidity to affect returns and then test our hypotheses in-sample and out-of-sample through portfolio sorts.

5.1. Two Channels for Return Predictability

There are two known channels through which LOB liquidity can affect returns. First, Cao et al. (2009) and Brogaard et al. (2014) argue that an imbalance between sell- and buy-side liquidity can result in short-term price pressure and document a positive relationship between Level I order imbalance and next-day stock returns. Translating their finding to the LOB, short-run returns should be high when the bid side is more liquid than the ask side ($MCI_B < MCI_A$). Second, the existence of a liquidity premium documented in numerous studies including Amihud and Mendelson (1986), Brennan and Subrahmanyam (1996), and Amihud (2002) implies that long-run returns increase in both lagged MCI_A and MCI_B.

Taken together, we expect the following. The first channel implies a positive relationship between imbalance (MCI_{IMB}) and returns, which is more pronounced for daily than for monthly returns. Both channels imply a positive relationship between ask illiquidity (MCI_A) and both daily and monthly returns. In contrast, the relationship between MCI_B and returns is less straightforward to predict. We expect the liquidity premium to result in a positive relationship between lagged MCI_B and monthly returns and the relationship between MCI_B and next-day returns to be less positive than for MCI_A. To test these predictions, we report results for in-sample predictive regressions of daily and monthly returns on our three MCI measures, as well as for daily and monthly portfolio sorts.

5.2. In-Sample Predictability

We first estimate linear regressions for the pooled sample. Specifically, we regress daily or monthly return on lagged values of MCI_A, MCI_B or MCI_{IMB} and control for lagged returns in each regression. Table 9 presents coefficient estimates and HAC standard errors. The left side displays the power of bid- versus ask-side liquidity in forecasting daily returns. In line with our expectations, daily returns increase in lagged MCI_{IMB} and MCI_A and less so in MCI_B. Signs and levels of significance are identical in regressions with monthly returns, displayed on the right side of the table. However, the MCI_B coefficient is no longer lower than the MCI_A coefficient, which is consistent with the prediction that the liquidity premia channel dominates at longer horizons.

[Table 8 about here.]

We complement the in-sample results for the pooled sample with firm-by-firm predictive regressions that include the same variables as before. We require a stock to have at least 24 monthly or 63 daily observations to be included in this analysis. Figure 5 presents the (winsorized) distribution of the coefficient estimates on the MCI measures from the firm-by-firm predictive regressions. We complement the histogram with the descriptive statistics of the (unwinsorized) distribution in Table 9.

Overall, coefficient signs and significances are similar to those found for the pooled sample.[12] At the daily level, an increase in both MCI_A and MCI_B predicts, on average, higher returns. In line with our expectations, the evidence is stronger for MCI_A, which has a positive and significant coefficient estimate for 25% of the stocks in our sample, compared with 14% for MCI_B. This is also reflected in the coefficient estimates on lagged MCI_{IMB}, which, on average, positively predicts daily returns. At the monthly level, results are similar for the predictive relationship between and returns and MCI_A as well as MCI_B. However, the predictive relationship between MCI_{IMB} and monthly returns is less clear at the firm level. The coefficient estimates on lagged MCI_{IMB} are almost symmetrically distributed around zero, with almost an equal number of positive and negative coefficient estimates and only approximately 8% that are significant. This suggests that MCI_{IMB} does not possess much predictive power at longer horizons and is consistent with the prediction that the liquidity premia channel dominates at longer horizons.

[Figure 5 about here.]

[Table 9 about here.]

[12] The adjusted R^2 of the firm-by-firm regressions are very low and are not reported for brevity. This is not surprising, as it is well known that stock returns are notoriously difficult to forecast.

5.3. Portfolio Sorts

Finally, we verify the robustness of the in-sample evidence by conducting portfolio sorts. Table 10 reports equally and value-weighted returns (Panel A) and four-factor alphas (Panel B) of portfolios formed on a daily basis. At the end of each trading day, stocks are assigned to a portfolio according to the quintile of their MCI measure and held until the end of the next trading day, when new portfolios are defined. Of course, these portfolio strategies are very costly to implement given the ongoing rebalancing. Our results therefore do not indicate whether a strategy is profitable after trading costs. Rows 1-5 report returns and alphas of the respective portfolio. For alphas, reported in Panel B, we also display the level of significance to test whether they are different from zero. We omit the results of these tests in Panel A, as all portfolio returns are significantly positive. Table 11 reports the same results but for portfolios formed at the end of month M (depending on the average MCI value in M) and held until the end of month M+1.

[Table 10 about here.]

[Table 11 about here.]

Our results strongly support our initially stated hypotheses. Equally weighted daily returns and alphas increase significantly in MCI_A and MCI_{IMB} but not in MCI_B. Over one year (252 trading days), the return on the equally weighted high MCI_{IMB} portfolio is 21.17% higher than that of the low MCI_{IMB} portfolio. In fact, the four-factor alpha of the latter is negative, indicating that there is both upward and downward price pressure. The high MCI_{IMB} quintile alpha of 5.948 basis points per day is substantial. However, entering a long position in high MCI_{IMB} implies buying stocks with an illiquid ask (sell) side, which is particularly costly. Comparing equally weighted to value-weighted results, the price pressure hypothesis only appears to hold for smaller stocks. The return and alpha of the high value-weighted MCI_{IMB} portfolio do not differ significantly from those of the low MCI_{IMB} portfolio. This is not true for MCI_A, which may be because it also captures the second channel through which liquidity affects returns, liquidity premia. These become more visible in monthly portfolio sorts. In line with our hypothesis, returns and alphas increase in both MCI_A and MCI_B. While the difference between high and low MCI is not statistically significant, this is likely due to the relatively low number of months covered by our sample, which limits the power of this monthly analysis. In sum, our results are in line with the argument that lagged liquidity can affect short-run returns by creating price pressure and is positively related to long-run return because investors demand a premium for holding illiquid stocks.

6. Conclusion

Recent studies have highlighted the importance of separating sell- and buy-side liquidity. In this article, we use limit order book data to study bid- and ask-side liquidity at a daily frequency over the period 2002-2012. We contribute to the data on market microstructure, liquidity, and asset pricing in multiple ways. First, we introduce the marginal cost of immediacy as a new measure of Level II liquidity that exhausts the information on price and quantity in a given LOB. Our measure is conceptually similar to some liquidity measures obtained from Level I data such as Kyle's λ and Brennan et al.'s measures of buy- and sell-side liquidity. Second, to the best of our knowledge, we are the first to explore the LOB data provided by TRTH for a comprehensive sample of NYSE stocks. We describe how to clean the TRTH data and how to match them to the CRSP files. We compare liquidity estimates based on Level II data to common Level I liquidity estimates and examine the time-series evolution of bid- and ask-side liquidity. Fourth, we disentangle the determinants of bid- and ask-side liquidity at the aggregate level and at the firm level and provide evidence of persistent contrarian behavior in limit orders. In addition, we document that liquidity commonality varies substantially over time and differs between bid- and ask- side. Finally, we show that LOB liquidity and imbalance predict returns at the daily and monthly levels.

The relevance of our results extends beyond academic research. Our analysis of changes in bid- and ask-side liquidity around the 2008 short selling ban can help *policy makers* to understand the mechanisms through which regulatory interventions affect liquidity provision. We document that the ask- but not bid-side liquidity of financial stocks deteriorated during the ban. Furthermore, we show that the liquidity of non-financial firms also changed significantly, indicating potential spillover effects. Our study can help *investors* to understand when differentiating between ask- and bid-side liquidity matters. At times, it can be cheap to buy but expensive to sell and vice versa. Strategies intended to reduce transaction costs should depend on the trade direction. Furthermore, we document that LOB imbalances reflect price pressure and can predict daily returns, while the predictive power of LOB liquidity for monthly returns appears to be driven by liquidity premia.

Our results indicate potential for future research using limit order book data. For instance, it would be interesting to further examine the impact of regulatory intervention on liquidity provision by studying the effect of Regulation SHO on the LOB. Comparing the LOBs of NASDAQ and NYSE can provide insights into the relevance of differences in trading mechanisms on liquidity. Finally, future research could explore whether information in the LOB can be used to predict the higher moments of the return distribution.

Appendix

A-1. Matching TRTH and CRSP

The primary security identifier in the TRTH data is the Reuters Instrument Code (RIC), which consists of a security's ticker symbol and an optional extension indicating specific exchanges or security types. The primary identifier of a security in the CRSP database is its Permno. The only additional identifier in our TRTH dataset is the company name. As tickers and company names vary over time and across databases, they are known to be unreliable identifiers for matching databases.

We therefore establish a table allowing us to link TRTH and CRSP data on a daily basis using three matching criteria: the level of closing prices, the correlation between closing price returns and similarities in the ticker or company name. In sum, we consider the criteria to be strict and, accordingly, the resulting link to be conservative. Despite the strictness of the link, we nevertheless cover more than 98% of the volume in NYSE common stocks included in the CRSP database. The following describes the matching procedure in greater detail.

In a first step, we identify the CRSP security that best matches the price patterns of a TRTH security in a given month from the entire universe of NYSE-listed securities reported in CRSP for each TRTH stock-month. To do so, we need daily time series of closing prices. While a CRSP closing price is typically the last transaction price, we derive the TRTH closing price as the Level-1 midprice of the last book snapshot reported for a stock during trading hours. We then gather the time series of daily closing prices reported in TRTH and CRSP into monthly blocks. Stock-months with fewer than 10 prices are excluded from the matching procedure at this stage. Links for such stock-months with very few observations are established using the links obtained from the months around them in a later step.

We now compute two matching parameters, price distance and return correlation. The former is the average absolute difference between the time series of prices for a given stock in a given month. The second is the correlation between returns computed from the prices. If, in a given month, the CRSP security with the highest correlation is also the one with the lowest price distance, we save the CRSP Permno identifier together with the value of both matching parameters. If this is not the case, but the CRSP stock with the highest correlation with the TRTH stock in a given month has an average price distance of less than l%, we use this as the best match and again save both parameters. If no best match has been identified yet and the correlation computed for the security with the lowest price distance is at least 90%, we use the minimum-distance security as the best match.

This step produced a matrix containing a Permno for each RIC-month for which a best match was identified. We observe very few outliers in these Permno time series, indicating the validity of the employed matching criteria. We replace these outliers with the Permno link from the surrounding months.

In a next step, we validate the monthly link by comparing the RIC identifier – which contains the security's ticker symbol – to all CRSP tickers reported in the CRSP names file for a given Permno. If we find a CRSP ticker symbol that fully matches the RIC ticker or only differs by one- or two-digit ticker extensions included in the RIC, we consider a link valid. For all non-verified links, we manually compare the company names from CRSP and TRTH. If these are not identical, we discard the link. Out of 6959 TRTH securities for which we identify a link in at least one month, only 14 are not verified based on ticker or company name.

Finally, we translate the monthly linking table to a daily one to avoid losing data from stock-months with fewer than 10 price observations. In a month without a link, we use the link from the preceding or subsequent month.[13] We then identify the beginning (ending) day of a link as the first (last) one with a price distance of less than 5%.

[13]This never causes a conflict because there are no occurrences of stock-months without link that are preceded by a link that is different from the subsequent link.

A-2. Sample Selection

Securities in the TRTH database are identified by the Reuters Instrument Code (RIC). Between 2002 and May 2006, the MD database includes RICs ending with ".CX", <.CX> denoting Nasdaq's Computer-Assisted Execution System (CAES). According to Thomson Reuters, the CAES quote is the best aggregated quote from market makers participating in the Intermarket Trading System (ITS). We compared data entries between securities with RICs ending in .CX to their counterpart not ending in .CX for randomly selected dates and found that the data are fully identical. We therefore exclude all securities with RICs with a .CX ending.

After the exclusion of these double entries, our initial dataset includes 52.44 billion snapshots. We require each snapshot to give a full picture of the book, excluding any snapshot for which order price or volume data are missing for any level of the book. This reduces the number of snapshots to 49.61 billion. We next limit the sample to snapshots for which prices are strictly monotonically increasing over the 20 levels of the book. This eliminates all negative bid-ask spreads and decreases the sample size to 44.83 billion snapshots. In line with Korajczyk and Sadka (2008), we delete entries with very large bid-ask spreads and those not reported for NYSE trading hours.[14] We then delete observations with a midprice of $1 or less or $1000 or more, yielding 44.49 billion observations. Limiting our sample to securities for which we are able to establish a link to data reported in the Center for Research in Security Prices (CRSP) files as detailed in Appendix A-1 further decreases this number to 42.83 billion. Removing all securities other than ordinary common shares identified by a CRSP sharecode of 10 or 11 results in our sample of 31.32 billion LOB snapshots. Finally, deleting all firm-days with fewer than 100 snapshots in the TRTH data or fewer than 100 trades in the matched TAQ dataset as well as observations without return data in CRSP or balance sheet information in Compustat results in a sample of 31.19 billion observations reported for 2,103 stocks over

[14]We delete observations with a Level-1 (10) ask price that is more than 25% (250%) higher than the Level-1 (10) bid price.

2,740 trading days and 3.36 million stock-day observations. The following overview summarizes how the above criteria reduce the sample size.[15]

Criterium	**billion snapshots left**
Delete snapshots ...	
... with missing price or volume	49.61
... without strictly monotone price schedule	44.83
... with large bid-ask spreads or outside trading hours	44.49
... without link to CRSP data	42.83
... which are not ordinary common shares	31.32
... with fewer than 100 TAQ or TRTH observations	
per day or missing CRSP or Compustat data:	31.19

A-3. Comparison of the Marginal Cost of Immediacy and the Xetra Liquidity Measure

MCI is conceptually and empirically closely related to the Hypothetical Transaction Cost (HTC_V) for a given order volume V. HTC_V measures how much an investor pays in bid-ask spread if he buys and sells V in the book instantaneously and is published by the German Stock Exchange under the label "Xetra Liquidity Measure" to describe the LOB liquidity of traded securities.[16] In Figure 1, the $HTC_{V=USD\ 1mio}$ corresponds to the length of the horizontal line and equals approximately 5.5 basis points. Relative to HTC_V, MCI has two main advantages. First, MCI exhausts all information from the book, while HTC_V only uses information up to the level at which the cumulative volume reaches V.[17] Second, MCI can be calculated for any stock, while HTC_V is restricted to LOBs with a cumulative bid and ask volume of at

[15] Variables obtained from CRSP include share code, market capitalization (number of shares outstanding times closing price), daily returns, and Standard Industry Classification (SIC) codes. Compustat balance sheet information comprises the quarterly items of total assets, cash and short-term investments, total liabilities, net income, and equity book values.

[16] See Roesch and Kaserer (2014).

[17] In the example given in Figure 1, the measure ignores the information contained in levels 7-10 of the ask schedule.

least V. This makes HTC unsuitable for studying a wide cross-section of stocks. Although we accordingly do not report results for HTC, we observe that in a given size category (allowing the HTC measure to be meaningful across all stocks), HTC_V and MCI are almost perfectly correlated.

We aggregate all TRTH measures on a daily basis by computing their equally weighted average. To avoid spurious results due to extreme outliers, we winsorize all Level I and Level II variables by setting the value below the 0.1 fractile (above the 99.9th percentile) equal to the value of the 0.1 (99.9) fractile.

A-4. Calculation of Level I Measures

The following details the calculation of Level 1 measures of liquidity computed from TAQ data. The first set of level I liquidity measures is based on Holden and Jacobsen (2014). The authors compare the quality of a wide range of liquidity measures constructed from the expensive Daily TAQ (DTAQ) database to measures based on the more commonly used Monthly TAQ (MTAQ) database. While they conclude that DTAQ is the first best, they suggest an interpolated time technique that makes it possible to improve the quality of MTAQ measures. We use these improved MTAQ measures as a benchmark to assess the marginal information content of our Level II measures. Given the substantial correlation between some of their measures, we report results only for the following measures. Whenever necessary, we use the Lee and Ready (1991) algorithm to classify orders into buys and sells.

The percent quoted spread over a certain time interval s during a trading day is given by:

$$PrctQuotedSpread_s = \frac{A_s - B_s}{M_s} \tag{7}$$

where A_s and B_s are the National Best Ask and Bid prices during a certain time interval s, respectively, and M_s is the corresponding the midquote during the same time interval, i.e., the average of A_s and B_s. Then, a stock's daily percent quoted spread is calculated as the time-weighted average of $PrctQuotedSpread_s$ over all time intervals in a given trading day.

The percent effective spread for a given trade k is defined as:

$$EffectiveSpreadPercent_k = \frac{2D_k(P_k - M_k)}{M_k} \tag{8}$$

where D_k is a trade direction indicator based on the Lee and Ready algorithm and is +1 if the kth trade is a buy and -1 if the kth trade is a sell. P_k is the transaction price of the kth rate, and M_k is the midquote price assigned to the kth trade based on the time interpolation technique proposed by Holden and Jacobsen (2014). Then a stock's daily percent effective spread is weighted average of $EffectiveSpreadPercent_k$ over all trades during a trading day where the weights are the dollar volume of each trade.

The total dollar depth available at both the best bid and ask prices for a time interval s is given by :

$$TotalDepthDollar_s = B_s \times Q_s^B + A_s \times Q_s^A \tag{9}$$

where Q_s^A and Q_s^B are the total quantities available at the best ask and bid quotes, respectively. A stock's daily total dollar depth can then be calculated as the time-weighted average of total dollar depth available over all time intervals. The total dollar depth imbalance between bid and ask sides is defined as the time-weighted average of the ratio of dollar depth available at the best ask price to total dollar depth available at both the best bid and ask prices.

The percent price impact for the kth trade is given by:

$$PercentPriceImpact_k = \frac{2D_k(M_{k+5} - M_k)}{M_k} \tag{10}$$

where M_{k+5} is the midquote price five minutes after the kth trade. The percent price impact for a stock in a given trading is the dollar-volume-weighted average of percent price impact over all trades.

We include two different measures of imbalance against which we can benchmark our principal measure of imbalance MCI_{IMB}. Similar to Holden and Jacobsen (2014), we compute a measure of order imbalance as:

$$OrderImbalance = \frac{N_B - N_S}{N_B + N_S}, \tag{11}$$

where N_B and N_S are the number of buys and sells according to the Lee and Ready (1991) trade classification. In addition, we calculate the Level 1 quantity imbalance as:

$$TotalDepthDollarImbalance = \frac{A_s \times Q_s^A}{B_s \times Q_s^B + A_s \times Q_s^A}, \tag{12}$$

where the denominator corresponds to Equation 9.

We estimate the Brennan et al. (2012) buy and sell lambdas following using the regression specification:

$$\Delta p_t = \alpha + \lambda_{buy}(q_t|q_t > 0) + \lambda_{sell}(q_t|q_t < 0) + \psi(D_t - D_{t-1}) + y_t, \tag{13}$$

where δp_t is the change in the price of a stock from one transaction to the next, q_t is the transaction quantity, $q_t > 0$ and $q_t < 0$ indicate whether a transaction is classified as a buy or sell according to the Lee and Ready (1991) algorithm and D_t is the sign of the order (+1 for buys, -1 for sells) at time t. Lambdas capture liquidity as price impact based on the relationship between order flows and returns and are conceptually derived from the work of Glosten and Harris (1988). According to their framework, y_t is the unobservable innovation in the fundamental value that, in the estimation, corresponds to the error term. ψ is the zero-quantity bid-ask spread. We divert from Brennan et al. (2012) in two ways. First, while Brennan et al. (2012) estimate the regression using one month of TAQ intraday data, we do so on a daily basis for consistency with our dataset. While this of course renders the estimates more noisy, they can still be used to obtain meaningful empirical results, as will be shown subsequently. Second, we define q as the dollar quantity traded rather than the number of shares traded. In sum, the Brennan et al. (2012) approach is an alternative for obtaining buy- and sell-side liquidity from widely available TAQ data.

The Amihud illiquidity measure for a trading day is defined as:

$$AmihudIlliq = \frac{1}{K}\sum_{k=1}^{K}\frac{|r_k|}{P_k * Q_k} \tag{14}$$

where r_k is the return on a stock between the $(k-1)^{th}$ and k^{th} trade, Q_k is the share volume corresponding to the kth trade, and K is the total number of trades in the trading day. The realized volatility of a stock for a given trading day is calculated as the sum of squared returns times $\sqrt{(}252)$, i.e.:

$$RealizedVol = \sqrt{252}\sum_{k=1}^{K}(r_k)^2 \tag{15}$$

We aggregate all TAQ liquidity measures that can be computed tick-by-tick to daily averages. Variables estimated from time series (*RealizedVol*, *AmihudIlliq*, and lambdas) are computed at daily frequency. To avoid spurious results due to extreme outliers, we winsorize all Level I and Level II variables by setting the value below the 0.1 fractile (above the 99.9th percentile) equal to the value of the 0.1 (99.9) fractile.

A-5. Calculation of PIN

PIN is the probability of informed trading described in Easley et al. (1996) and Easley et al. (2002). To obtain monthly estimates of PIN for each firm, we follow the approach described in Lin and Ke (2011). Specifically, we estimate the parameters of the structural model by maximizing the joint probability density function of the observed numbers of buys (B) and sells (S) in a given month for a given firm, rewritten in a more accurate fashion to reduce biases due to numerical problems:

$$\begin{aligned}\mathcal{L}(\theta|B,S) &= \log(\alpha\delta\exp(e_1 - e_{max}) + \alpha(1-\delta)\exp(e_2 - e_{max}) + (1-\alpha)\exp(e_3 - e_{max})) \\ &+ B\log(\varepsilon_b + \mu) + S\log(\varepsilon_s + \mu) - (\varepsilon_b + \varepsilon_s) + e_{max} - log(S!B!)\end{aligned}$$

where α is the probability of an information event, δ is the probability of a bad signal, ε_b and ε_s are the arrival rates of uninformed buys and sells,respectively, μ is the arrival rate of informed trades, $e_1 = -\mu - B\log(1+\mu/\varepsilon_b)$, $e_2 = -\mu - S\log(1+\mu/\varepsilon_s)$, $e_3 = -B\log(1+\mu/\varepsilon_b) - S\log(1+\mu/\varepsilon_s)$ and $e_max = \max(e_1, e_2, e_3)$. We restrict ε_b, ε_s and μ to be greater than 0.5 and both α and δ to be between 0.025 and 0.975. We use the NLP procedure (NonLinear Programming) of SAS to for the estimation. The PIN, which is the ratio of mean informed trades to mean total trades, can then be calculated based on the estimated parameters as follows:

$$\text{PIN} = \frac{\hat{\alpha}\hat{\mu}}{\hat{\alpha}\hat{\mu} + \hat{\varepsilon}_b + \hat{\varepsilon}_s} \tag{16}$$

References

Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets 5 (1), 31–56.

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid ask spread. Journal of Financial Economics 17 (2), 223–249.

Amihud, Y., Mendelson, H., Pedersen, L. H., 2012. Market liquidity: Asset pricing, risk, and crises. Cambridge University Press.

Aragon, G. O., Strahan, P. E., 2012. Hedge funds as liquidity providers: Evidence from the lehman bankruptcy. Journal of Financial Economics 103 (3), 570–587.

Beber, A., Pagano, M., 2013. Short-selling bans around the world: Evidence from the 2007-09 crisis. Journal of Finance 68 (1), 343–381.

Boehmer, E., Jones, C. M., Zhang, X., 2015. Potential pilot problems: Treatment spillovers in financial regulatory experiments. Unpublished Working Paper., 1–36.

Bollen, N. P., Smith, T., Whaley, R. E., 2004. Modeling the bid/ask spread: measuring the inventory-holding premium. Journal of Financial Economics 72 (1), 97–141.

Brennan, M., Huh, S.-W., Subrahmanyam, A., 2013. An analysis of the amihud illiquidity premium. Review of Asset Pricing Studies 3 (1), 133–176.

Brennan, M. J., Chordia, T., Subrahmanyam, A., Tong, Q., 2012. Sell-order liquidity and the cross-section of expected stock returns. Journal of Financial Economics 105 (3), 523–541.

Brennan, M. J., Subrahmanyam, A., 1996. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. Journal of Financial Economics 41 (3), 441–464.

Brogaard, J., Hendershott, T., Riordan, R., 2014. High-frequency trading and price discovery. Review of Financial Studies 27 (8), 2267–2306.

Brunnermeier, M. K., Pedersen, L. H., 2009. Market liquidity and funding liquidity. Review of Financial studies 22 (6), 2201–2238.

Campbell, J. Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. Journal of Finance 63 (6), 2899–2939.

Cao, C., Hansch, O., Wang, X., 2009. The information content of an open limit-order book. Journal of Futures Markets 29 (1), 16–41.

Carhart, M. M., 1997. On persistence in mutual fund performance. The Journal of finance 52 (1), 57–82.

Chacko, G. C., Jurek, J. W., Stafford, E., 2008. The price of immediacy. The Journal of Finance 63 (3), 1253–1290.

Choi, J. Y., D. Salandro, D., Shastri, K., 1988. On the estimation of bid-ask spreads: Theory and evidence. Journal of Financial and Quantitative Analysis 23 (2), 219–230.

Chordia, T., Roll, R., Subrahmanyam, A., 2000. Commonality in liquidity. Journal of Financial Economics 56 (1), 03–28.

Chordia, T., Roll, R., Subrahmanyam, A., 2001. Market liquidity and trading activity. Journal of Finance 56, 501–530.

Chordia, T., Roll, R., Subrahmanyam, A., 2002. Order imbalance, liquidity, and market returns. Journal of Financial Economics 65 (1), 111–130.

Comerton-Forde, C., Hendershott, T., Jones, C. M., Moulton, P. C., Seasholes, M. S., 2010. Time variation in liquidity: The role of market-maker inventories and revenues. Journal of Finance 65 (1), 295–331.

Coughenour, J. F., Saad, M. M., 2004. Common market makers and commonality in liquidity. Journal of Financial Economics 73 (1), 37–69.

Easley, D., Hvidkjaer, S., O'hara, M., 2002. Is information risk a determinant of asset returns? The Journal of Finance 57 (5), 2185–2221.

Easley, D., Kiefer, N. M., O'Hara, M., Paperman, J. B., 1996. Liquidity, information, and infrequently traded stocks. Journal of Finance 51, 1405–1436.

Feldhuetter, P., 2011. The same bond at different prices: Identifying search frictions and selling pressures. Review of Financial Studies 25 (4), 1155–1206.

Fong, K., Holden, C., Trzcinka, C., 2014. What are the best proxies for global research? Unpublished working paper. Indiana University.

Garman, M. B., 1976. Market microstructure. Journal of Financial Economics 3 (3), 257–275.

Glosten, L. R., Harris, L. E., 1988. Estimating the components of the bid/ask spread. Journal of Financial Economics 21 (1), 123–142.

Glosten, L. R., Milgrom, P. R., 1985. Bid, ask, and transaction prices in a specialist market with heterogeneously informed agents. Journal of Financial Economics 14 (1), 71–100.

Grossman, S. J., Miller, M. H., 1988. Liquidity and market structure. The Journal of Finance 43 (3), 617–633.

Hameed, A., Kang, W., Viswanathan, S., 2010. Stock market declines and liquidity. The Journal of Finance 65 (1), 257–293.

Hasbrouck, J., 2004. Liquidity in the futures pits: Inferring market dynamics from incomplete data. Journal of Financial and Quantitative Analysis 39 (3), 305–326.

Ho, T. S. Y., Macris, R. H., 1984. Dealer bid-ask quotes and transaction prices: An empirical study of some amex options. Journal of Finance 39 (1), 23–45.

Ho, T. S. Y., Stoll, H. R., 1981. Optimal dealer pricing under transactions and return uncertainty. Journal of Financial Economics

9 (1), 47–73.

Holden, C., 2009. New low-frequency liquidity measures. Journal of Financial Markets 12 (4), 778–813.

Holden, C. W., Jacobsen, S., 2014. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. Journal of Finance 69 (4), 1747–1785.

Holden, C. W., Jacobsen, S. E., Subrahmanyam, A., 2014. The empirical analysis of liquidity. Kelley School of Business Research Paper (2014-09).

Huang, R. D., Stoll, H. R., 1996. Dealer versus auction markets: A paired comparison of execution costs on nasdaq and the nyse. Journal of Financial Economics 41 (3), 313–357.

Kalay, A., Wohl, A., 2009. Detecting liquidity traders. Journal of Financial and Quantitative Analysis 44 (1), 29–54.

Kamara, A., Lou, X., Sadka, R., 2008. The divergence of liquidity commonality in the cross-section of stocks. Journal of Financial Economics 89 (3), 444–466.

Karolyi, G. A., Lee, K.-H., Van Dijk, M. A., 2012. Understanding commonality in liquidity around the world. Journal of Financial Economics 105 (1), 82–112.

Korajczyk, R. A., Sadka, R., 2008. Pricing the commonality across alternative measures of liquidity. Journal of Financial Economics 87 (1), 45–72.

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

Lee, C., Ready, M., 1991. Inferring trade direction from intraday data. Journal of Finance 46 (2), 733–746.

Lesmond, D. A., Ogden, J. P., Trzcinka, C. A., 1999. A new estimate of transaction costs. Review of Financial Studies 12 (5), 1113–1141.

Lin, H.-W. W., Ke, W.-C., 2011. A computing bias in estimating the probability of informed trading. Journal of Financial Markets 14, 625–640.

Marshall, B. R., Nguyen, N. H., Visaltanachoti, N., 2011. Commodity liquidity measurement and transaction costs. Review of Financial Studies, 1–40.

Pastor, L., Stambaugh, R. F., 2003. Liquidity risk and price discovery. Journal of Political Economy 111 (3), 642–685.

Petersen, M. A., 2009. Estimating standard errors in finance panel data sets: Comparing approaches. Review of Financial Studies 22 (1), 435–480.

Roesch, C. G., Kaserer, C., 2014. Reprint of: Market liquidity in the financial crisis: The role of liquidity commonality and flight-to-quality. Journal of Banking and Finance 45 (1), 152–170.

Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. Journal of Finance 39 (4), 1127–1139.

Rosu, I., 2009. A Dynamic Model of the Limit Order Book. The Review of Financial Studies 22 (11), 4601–4641.

Stoll, H. R., 1978. The supply of dealer services in securities markets. Journal of Finance 33 (4), 1133–1151.

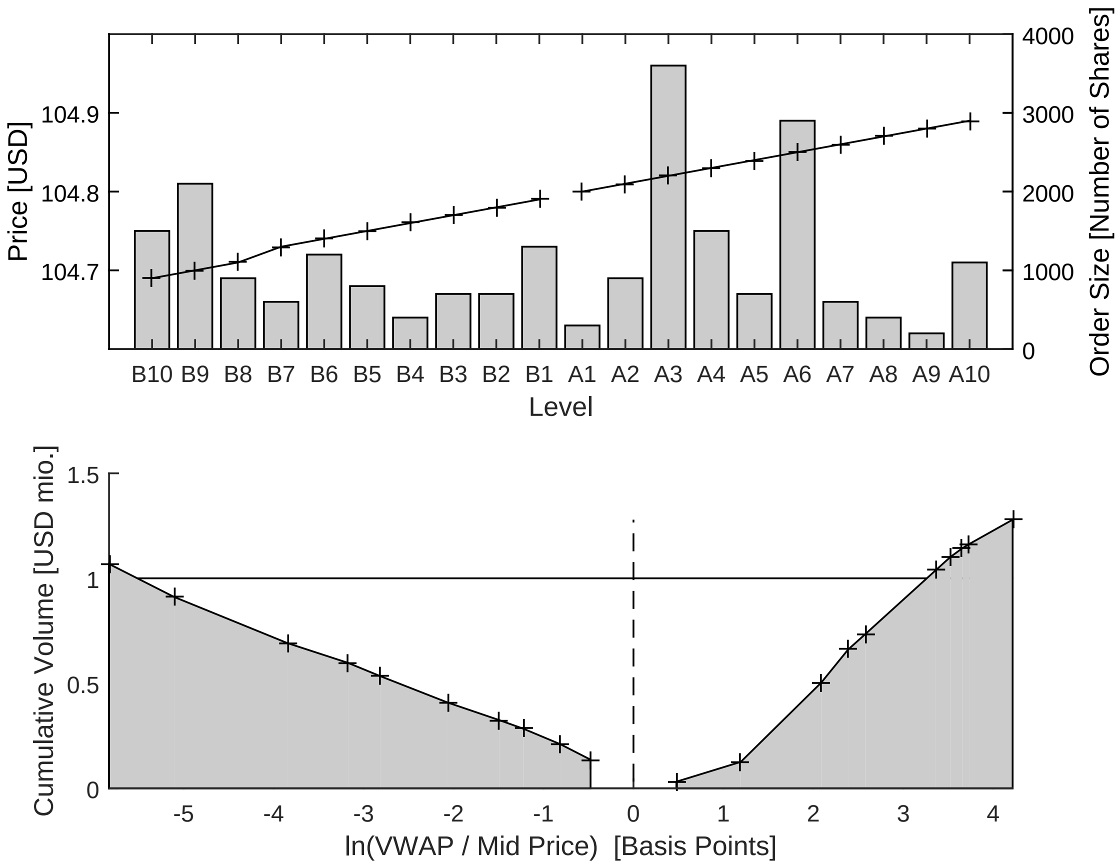


Figure 1: *A sample snapshot of a limit order book*

This figure displays the snapshot of the limit order book for IBM on July 2nd, 2007 at 3 PM. The upper graph plots the limit order prices (line measured on the left axis) and quantities (bars measured on the right axis) as reported. The left and right sides show bid (B1 to B10) and ask (A1 to A10) information for the first ten levels of the book, respectively. The lower panel presents a transformation of the reported limit order book, where the logarithm of the volume-weighted average price (VWAP) of the ask (bid) side scaled by the midprice measured in basis points on the x-axis is plotted against the cumulative dollar volume at the first ten levels of the ask (bid) side measured in USD million on the y-axis. Each plus sign (+) corresponds to a level of the limit order book. The midprice is denoted by zero on the x-axis, and the bid side is measured to the left of the midprice, while the ask side is on the right.

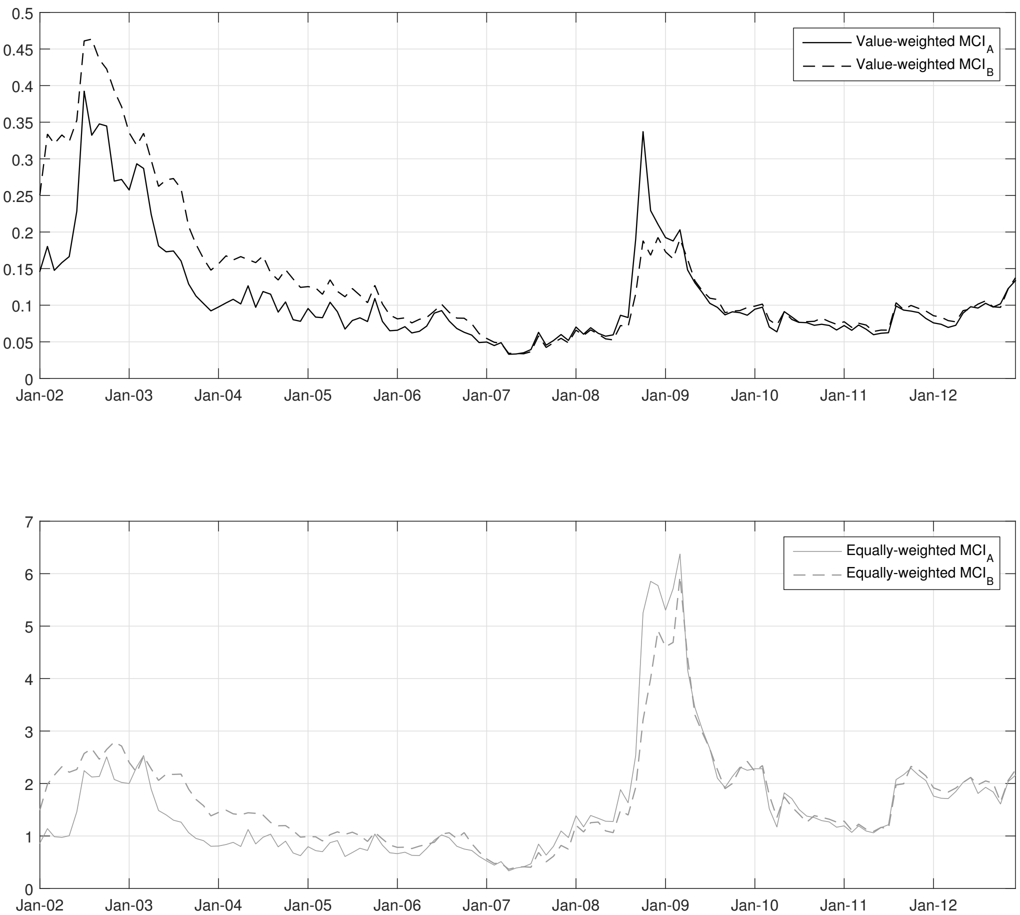


Figure 2: *Monthly time series of ask- and bid-side transaction costs*
This figure displays the monthly averages of daily market-wide ask- and bid-side transaction costs MCI_A and MCI_B in basis points, computed as the value-weighted (upper panel) and equally weighted (lower panel) averages of ask- and bid-side transaction costs for individual stocks.

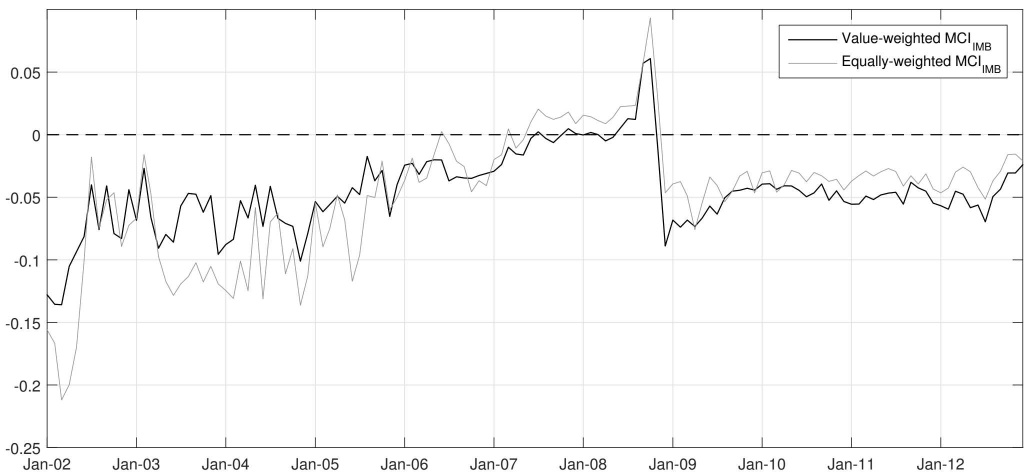


Figure 3: *Monthly time series of MCI imbalances*

This figure displays the value-weighted (black line) and equally weighted (grey line) averages of imbalances between ask- and bid-side transaction costs $MCI_{IMB} = (MCI_A - MCI_B)/(MCI_A + MCI_B)$ of individual stocks. For visualization purposes, we display monthly averages of daily data.

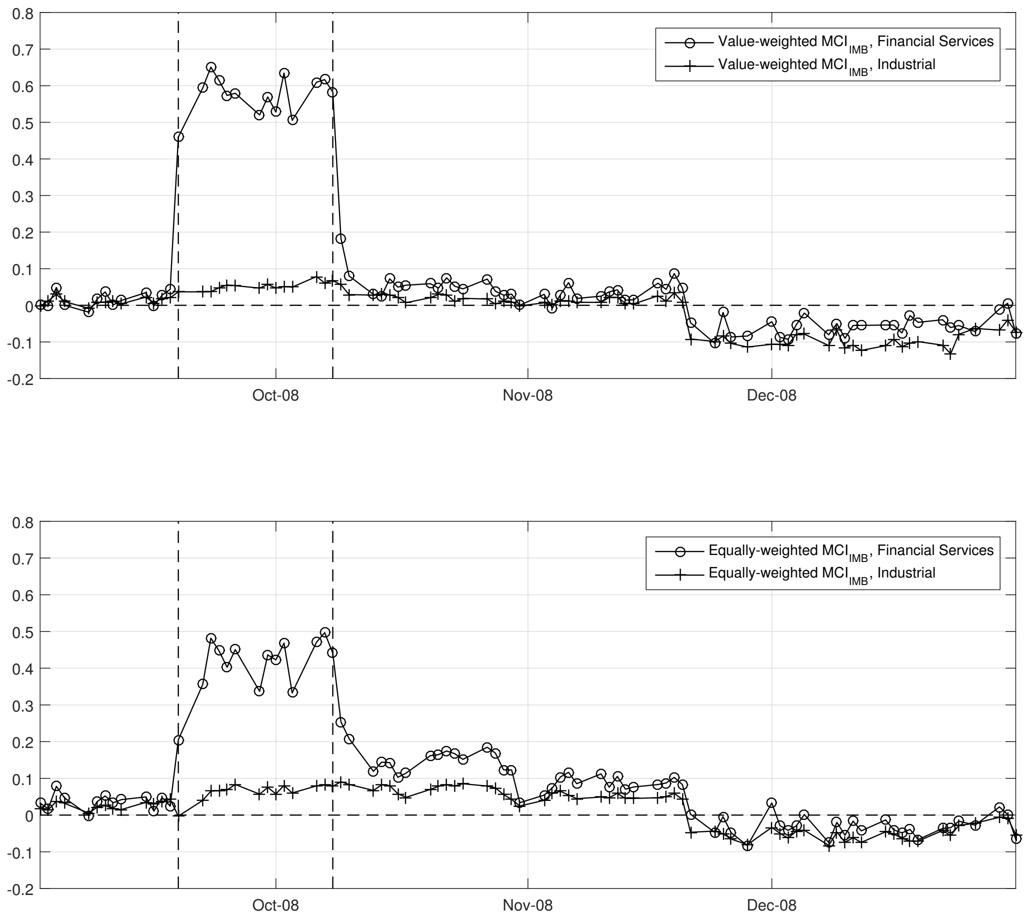
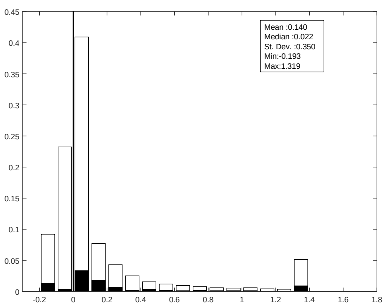
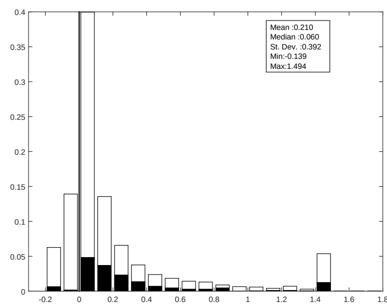


Figure 4: *Daily time series of transaction cost imbalances around the 2008 short selling ban*
This figure displays value-weighted (upper panel) and equally weighted (lower panel) averages of daily imbalances between ask- and bid-side transaction costs $MCI_{IMB} = (MCI_A - MCI_B)/(MCI_A + MCI_B)$ around the short selling ban imposed on financial stocks (SIC code 6000-6999) between September 19th and October 8th, 2008. *o* and + denote financial and non-financial stocks, respectively. The vertical dashed lines denote the beginning and end of the short selling ban.

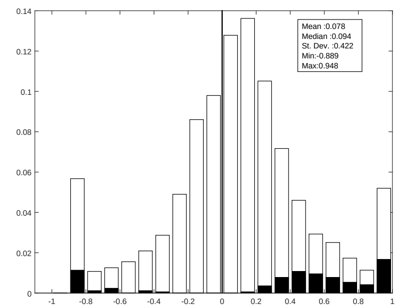
Histogram of predictive coefficients for daily returns



(a) MCI_B

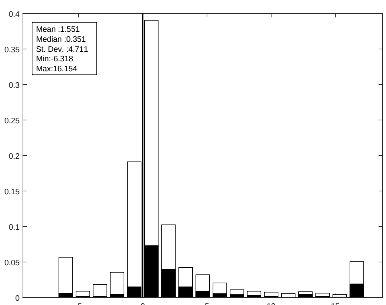


(b) MCI_A

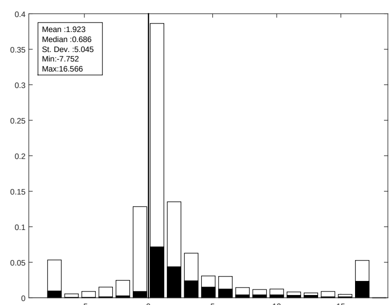


(c) MCI_{IMB}

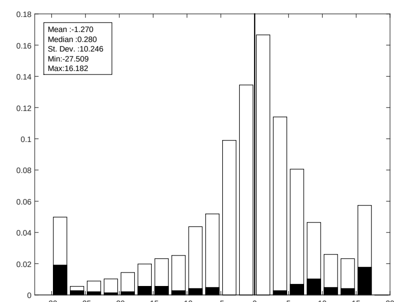
Histogram of predictive coefficients for monthly returns



(d) MCI_B



(e) MCI_A



(f) MCI_{IMB}

Figure 5: *Histogram of firm-by-firm predictive coefficients*
This figure presents the cross-sectional distribution of the coefficient estimates on the lagged values of MCI_B (left column), MCI_A (middle column) and MCI_{IMB} (right column) from the predictive regressions of daily (upper panel) and monthly (lower panel) stock returns on their own lagged values and each of these variables, separately. The histograms and summary statistics are based on the distributions winsorized at 5% and 95%. The regressions are estimated via least squares separately for each stock with at least 63 observations for daily and 24 for monthly analysis. The white bars present the percentage of coefficient estimates that fall in each bin, while the black bars present the percentage of coefficient estimates that are significant at the 5% level based on HAC standard errors. The vertical line is at zero and distinguishes positive from negative coefficient estimates.

Table 1: *Descriptive statistics.*

This table presents descriptive statistics for variables computed from TRTH Level II LOB data (Panel (A)) and TAQ Level I data (Panel (B)). The former include the relative (as opposed to dollar) bid-ask spread for level 1 and level 10 ($PrctQuotedSpread_{L=1}$ and $PrctQuotedSpread_{L=10}$), the dollar volume available at the first, as well as all ten observed levels of the LOB ($Vlm_{L=1}$ and $Vlm_{L=10}$), the imbalances between ask and bid side volume ($Vlm_{IMB,L=1}$ and $Vlm_{IMB,L=10}$), the bid- and ask-side marginal cost of immediacy and their logged values (MCI_B, MCI_A, $logMCI_A$, $logMCI_B$) and its bid-ask imbalance (MCI_{IMB}), together with the number of book snapshots observed per firm-day. Measures computed for TAQ data include the relative best bid-ask spread (*PrctQuotedSpread*), the relative effective spread (*EffectiveSpreadPercent*), the total volume available at the best bid and ask ($TotalDepthDollar$) and the according bid-ask imbalance ($TotalDepthDollarImbalance$), and the relative price impact ($PercentPriceImpact$) as used by Holden and Jacobsen (2014), as well as the annualized realized volatility ($RealizedVol$), the Amihud (2002) measure of illiquidity ($AmihudIlliq$), the order imbalance (buys - sells)/(buys + sells) ($OrderImbalance$), and the Brennan et al. (2012) buy and sell lambdas and their imbalance ($SellLambda$, $BuyLambda$, $LambdaImbalance$). The calculation of the effective spread, the relative price impact, order imbalance, and lambdas require the classification of trades into buys and sells, which we obtain using the Lee and Ready (1991) algorithm. bp indicates basis points.

(a) TRTH Variables

	$Mean$	$StdDev$	1^{st}	25^{th}	50^{th}	75^{th}	99^{th}
$PrctQuotedSpread_{L=1}$ [bp]	21.442	26.938	2.652	7.305	12.808	24.185	139.615
$PrctQuotedSpread_{L=10}$ [bp]	667.240	1,243.459	29.210	84.025	184.905	622.924	6,061.729
$Vlm_{L=1}$ [USD mio.]	50.688	81.849	2.625	16.917	31.039	57.381	330.817
$Vlm_{L=10}$ [USD mio.]	913.882	1,072.418	49.021	306.952	614.791	1,118.703	5,410.087
$Vlm_{IMB,L=1}$	0.025	0.104	−0.236	−0.034	0.017	0.078	0.324
$Vlm_{IMB,L=10}$	0.054	0.171	−0.431	−0.019	0.039	0.109	0.622
MCI_B [bp per USD 1,000]	1.643	6.144	0.007	0.067	0.224	0.896	25.976
MCI_A [bp per USD 1,000]	1.544	6.482	0.006	0.060	0.193	0.746	26.311
$logMCI_B$	−1.352	1.850	−5.031	−2.700	−1.496	−0.109	3.257
$logMCI_A$	−1.488	1.831	−5.075	−2.816	−1.645	−0.294	3.270
MCI_{IMB}	−0.044	0.211	−0.713	−0.113	−0.036	0.035	0.590
Number of LOB snapshots ['000s]	9.267	6.308	0.216	4.088	8.359	13.763	24.224
Firm-day Observations							3,365,528

(b) TAQ Variables

	$Mean$	$StdDev$	1^{st}	25^{th}	50^{th}	75^{th}	99^{th}
PrctQuotedSpread [bp]	15.839	17.505	2.134	6.076	10.444	18.623	90.814
EffectiveSpreadPercent [bp]	14.482	18.756	2.157	5.472	9.025	16.038	92.861
$TotalDepthDollar$ [USD mio.]	1.031	2.698	0.057	0.269	0.483	0.922	10.186
$TotalDepthDollarImbalance$ [%]	0.524	0.100	0.255	0.470	0.516	0.575	0.802
$PercentPriceImpact$ [bp]	8.407	14.585	−11.705	2.044	4.925	10.366	66.249
$RealizedVol$ [%]	35.276	25.942	8.606	19.570	28.050	41.909	138.791
$AmihudIlliq$	0.322	1.082	0.006	0.028	0.067	0.190	5.114
$OrderImbalance$	0.036	0.133	−0.307	−0.040	0.027	0.109	0.400
$Sell\ Lambda$	0.383	1.307	0.000	0.015	0.065	0.252	5.483
$Buy\ Lambda$	0.368	1.379	0.000	0.011	0.052	0.215	5.552
$Lambda\ Imbalance$	−0.378	44.554	−71.126	−2.762	−1.152	2.322	69.960
Number of trades ['000s]	6.656	17.483	0.132	0.785	1.957	5.698	68.370
Firm-day Observations							3,365,528

Table 2: *Correlations.*

This table reports correlations between TRTH and TAQ liquidity measures for the pooled sample. The former include the relative (as opposed to dollar) bid-ask spread for level 1 and level 10 ($PrctQuotedSpread_{L=1}$ and $PrctQuotedSpread_{L=10}$), the dollar volume available at the first, as well as all ten observed levels of the LOB ($Vlm_{L=1}$ and $Vlm_{L=10}$), the imbalances between ask and bid side volume ($Vlm_{IMB,L=1}$ and $Vlm_{IMB,L=10}$), the bid- and ask-side marginal cost of immediacy and their logged values (MCI_B, MCI_A, $logMCI_A$, $logMCI_B$) and its bid-ask imbalance (MCI_{IMB}). Measures computed for TAQ data include the relative best bid-ask spread (*PrctQuotedSpread*), the relative effective spread (*EffectiveSpreadPercent*), the total volume available at the best bid and ask (*TotalDepthDollar*) and the according bid-ask imbalance (*TotalDepthDollarImbalance*), and the relative price impact (*PercentPriceImpact*) as used by Holden and Jacobsen (2014), as well as the annualized realized volatility (*RealizedVol*), the Amihud (2002) measure of illiquidity (*AmihudIlliq*), the order imbalance (buys - sells)/(buys + sells) (*OrderImbalance*), and the Brennan et al. (2012) buy and sell lambdas and their imbalance (*SellLambda*, *BuyLambda*, *LambdaImbalance*). The calculation of the effective spread, the relative price impact, order imbalance, and lambdas require the classification of trades into buys and sells, which we obtain using the Lee and Ready (1991) algorithm.

(a) TRTH Variables

	1	2	3	4	5	6	7	8	9	10	11
(1) $PrctQuotedSpread_{L=1}$ [bp]											
(2) $PrctQuotedSpread_{L=10}$ [bp]	0.55										
(3) $Vlm_{L=1}$ [USD mio.]	−0.21	−0.12									
(4) $Vlm_{L=10}$ [USD mio.]	−0.30	−0.16	0.76								
(5) $Vlm_{IMB,L=1}$	−0.03	0.08	0.05	0.01							
(6) $Vlm_{IMB,L=10}$	0.06	0.19	−0.05	−0.01	0.23						
(7) MCI_B [bp per USD 1,000]	0.67	0.49	−0.12	−0.17	0.02	0.15					
(8) MCI_A [bp per USD 1,000]	0.65	0.49	−0.11	−0.16	−0.05	−0.04	0.79				
(9) $logMCI_B$	0.70	0.60	−0.42	−0.61	0.06	0.22	0.52	0.44			
(10) $logMCI_A$	0.72	0.59	−0.43	−0.62	−0.01	0.00	0.48	0.50	0.95		
(11) MCI_{IMB}	0.03	−0.02	0.01	−0.02	−0.28	−0.70	−0.12	0.17	−0.16	0.14	

(b) TAQ Variables

	1	2	3	4	5	6	7	8	9	10	11
PrctQuotedSpread [bp]	0.90	0.54	−0.20	−0.31	−0.02	0.05	0.64	0.62	0.71	0.73	0.03
EffectiveSpreadPercent [bp]	0.78	0.47	−0.14	−0.22	0.00	0.04	0.56	0.55	0.58	0.60	0.02
TotalDepthDollar [USD mio.]	−0.13	−0.10	0.87	0.70	−0.02	−0.07	−0.08	−0.07	−0.34	−0.33	0.04
TotalDepthDollarImbalance [%]	−0.07	0.05	0.06	0.03	0.55	0.27	−0.02	−0.08	0.04	−0.05	−0.30
OrderImbalance	−0.05	0.11	0.03	0.00	0.25	0.16	−0.01	−0.05	0.06	0.01	−0.19
PercentPriceImpact [bp]	0.54	0.39	−0.09	−0.16	0.03	0.05	0.36	0.35	0.44	0.44	0.01
RealizedVol [bp]	0.58	0.25	−0.20	−0.23	−0.04	−0.01	0.36	0.38	0.43	0.46	0.08
AmihudIlliq	0.73	0.37	−0.11	−0.17	−0.04	0.00	0.74	0.75	0.47	0.49	0.06
Sell Lambda	0.57	0.31	−0.14	−0.18	−0.02	0.03	0.58	0.57	0.43	0.44	0.02
Buy Lambda	0.55	0.29	−0.13	−0.17	−0.05	0.00	0.56	0.59	0.40	0.42	0.06
Lambda Imbalance	0.00	0.00	0.00	0.00	−0.01	0.00	0.00	0.00	0.00	0.00	0.00

Table 3: *Changes in the LOB around short selling ban*
This table shows L10 spreads and volume before and after the start of the 2008 short selling ban for financial services firms (Panel A) and all other firms (Panel B). *a* and *b* indicate that the change from pre to post value is significant at the 1% and 5% levels, respectively.

(a) Firms in financial services industry: SIC code 6000-6999 (N_{pre} = 201, N_{post} = 198)

	Average		*Median*	
	Pre	*Post*	*Pre*	*Post*
Spread to Midprice (L10, Bid) [%]	1.244	1.218	0.538	0.817^{a}
Spread to Midprice (L10, Ask) [%]	1.431	5.053^{a}	0.562	3.311^{a}
$Vlm_{L=10,Bid}$ [mio.]	0.348	0.280	0.202	0.177
$Vlm_{L=10,Ask}$ [mio.]	0.334	0.281	0.190	0.183

(b) All other firms (N_{pre} = 1120, N_{post} = 1123)

	Average		*Median*	
	Pre	*Post*	*Pre*	*Post*
Spread to Midprice (L10, Bid) [%]	0.844	0.989^{b}	0.439	0.484^{b}
Spread to Midprice (L10, Ask) [%]	1.262	1.816^{a}	0.444	0.540^{a}
$Vlm_{L=10,Bid}$ [mio.]	0.329	0.297^{b}	0.201	0.202
$Vlm_{L=10,Ask}$ [mio.]	0.320	0.304	0.197	0.215

Table 4: *Firm-level determinants of LOB liquidity.*

Regression Estimates: The first four columns of this table present the estimates from linear regressions of MCI measures on explanatory variables. We estimate regressions for $logMCI_A$ and $logMCI_B$ jointly via seemingly unrelated regressions (SUR) and test for the difference in coefficients using a Wald test. In all three regressions, standard errors (in parenthesis) are clustered at the firm level following Petersen (2009). Explanatory variables include the log of the daily dollar trading volume ($Turnover$), realized intraday volatility ($RealizedVol$), one minus the Herfindahl index measuring trading dispersion across exchanges ($Competition$), the Easley et al. (1996) probability of informed trading estimated over the previous month (PIN), the Lee and Ready (1991) order imbalance (buys - sells)/(buys + sells) ($OrderImbalance$) and its absolute value ($AbsOrderImbalance$), today's return (Ret), yesterday's return ($RetLag$), the Campbell et al. (2008) long-run excess return until yesterday ($ExretAvg$), logged market capitalization ($Size$), the Campbell et al. (2008) measures of log stock price winsorized at \$15, market to book ratio (MB), leverage ($Lever$), profitability ($Profit$), and cash holdings ($Cash$), a dummy for financial services firms (FS), one for days during which the 2008 short selling ban was active (Ban), and their interaction. Standard errors are reported below the coefficient estimates. a, b and c denote significance at the 1%, 5% and 10% significance levels, respectively. *Economic Significance:* The three last columns of this table display the product of coefficient estimates and the standard deviation for each explanatory variable.

	Regression Estimates				**Economic Significance**		
	$logMCI_A$	$logMCI_B$	$Diff$	MCI_{IMB}	$logMCI_A$	$logMCI_B$	MCI_{IMB}
Traditional Determinants							
$Turnover$	-0.563^a	-0.586^a	0.023^a	0.005^a	−0.979	−1.019	0.009
	(0.008)	(0.007)		(0.001)			
$RealizedVol$	1.616^a	1.611^a	0.005	-0.005^c	0.419	0.418	−0.001
	(0.023)	(0.022)		(0.003)			
$Competition$	-0.747^a	-0.996^a	0.249^a	0.041^a	−0.201	−0.269	0.011
	(0.018)	(0.018)		(0.003)			
PIN	0.147^a	0.128^a	0.019	−0.006	0.009	0.008	0.000
	(0.028)	(0.030)		(0.005)			
$AbsOrderImbalance$	0.082^a	0.198^a	-0.116^a	-0.057^a	0.008	0.018	−0.005
	(0.020)	(0.023)		(0.004)			
Trade Direction							
$OrderImbalance$	-0.092^a	0.401^a	-0.493^a	-0.219^a	−0.012	0.053	−0.029
	(0.014)	(0.017)		(0.003)			
Ret	-0.688^a	0.356^a	-1.043^a	-0.335^a	−0.021	0.011	−0.010
	(0.017)	(0.019)		(0.009)			
$RetLag$	-0.261^a	0.610^a	-0.870^a	-0.462^a	−0.008	0.019	−0.014
	(0.024)	(0.025)		(0.010)			
$ExretAvg$	-0.595^a	1.946^a	-2.542^a	-0.850^a	−0.023	0.075	−0.033
	(0.084)	(0.086)		(0.021)			
Control Variables							
$Size$	-0.321^a	-0.299^a	-0.023^a	-0.008^a	−0.488	−0.455	−0.012
	(0.010)	(0.009)		(0.001)			
PRC	-0.342^a	-0.357^a	0.015^b	0.005^b	−0.138	−0.144	0.002
	(0.018)	(0.018)		(0.003)			
MB	−0.005	-0.010^b	0.006^a	0.002^a	−0.006	−0.012	0.002
	(0.004)	(0.004)		(0.001)			
$Lever$	-0.098^a	-0.052^c	-0.046^a	-0.017^a	−0.023	−0.012	−0.004
	(0.028)	(0.028)		(0.005)			
$Profit$	-3.810^a	-3.217^a	-0.594^b	-0.179^b	−0.034	−0.029	−0.002
	(0.496)	(0.489)		(0.089)			
$Cash$	0.114	0.193^b	−0.079	-0.046^a	0.007	0.012	−0.003
	(0.081)	(0.078)		(0.013)			
FS	0.039^b	0.031^b	0.008	0.005^b	0.014	0.011	0.002
	(0.016)	(0.015)		(0.003)			
Ban	0.338^a	0.155^a	0.183^a	0.073^a	0.024	0.011	0.005
	(0.015)	(0.010)		(0.004)			
$Ban \times FS$	1.217^a	0.100^a	1.117^a	0.382^a	0.033	0.003	0.010
	(0.067)	(0.036)		(0.020)			
($Intercept$)	13.243^a	13.563^a	-0.320^a	-0.025^b			
	(0.084)	(0.080)		(0.010)			
$Adj.\ R^2$	0.849	0.835		0.075			

Table 5: *Time-series determinants of LOB liquidity.*
Panel A of this table presents estimates from linear regressions of equally weighted (EW) and value-weighted (VW) MCI measures on explanatory variables. We estimate regressions for average MCI_A and MCI_B jointly via seemingly unrelated regressions (SUR) and test for the difference in coefficients using a Wald test. Explanatory variables include the log of the market wide daily dollar trading volume ($Turnover_M$), the VIX volatility index (VIX), the ratio of stocks with positive to that with negative returns (*UpDown Ratio*), today's market return (Ret_M), yesterday's market return ($Ret_M Lag$), the TED spread as a measure of funding liquidity (*TED*), and a dummy equal to one for days during which the 2008 short selling ban was active (*Ban*). Standard errors are reported in parentheses. *a*, *b* and *c* denote significance at the 1%, 5% and 10% significance levels, respectively. Panel B of this table displays the product of coefficient estimates and the standard deviation for each explanatory variable.

(a) Regression Estimates

	EW				*VW*			
	MCI_A	MCI_B	$Diff.$	MCI_{IMB}	MCI_A	MCI_B	$Diff.$	MCI_{IMB}
Traditional Determinants								
$Turnover_M$	-1.272^a	-3.397^a	2.125^a	0.437^a	-0.499^a	-0.844^a	0.345^a	0.204^a
	(0.264)	(0.235)		(0.016)	(0.021)	(0.034)		(0.012)
VIX	11.668^a	9.550^a	2.118^a	0.031^a	0.568^a	0.549^a	0.019^b	-0.085^a
	(0.279)	(0.243)		(0.010)	(0.017)	(0.020)		(0.008)
Trade Direction								
UpDown Ratio	0.012	0.028	−0.016	-0.006^a	-0.011^a	-0.014^a	0.003^b	-0.002^b
	(0.027)	(0.025)		(0.001)	(0.002)	(0.002)		(0.001)
Ret_M	6.942^b	8.642^a	−1.700	-0.260^a	0.893^a	1.184^a	-0.291^b	-0.823^a
	(2.884)	(2.418)		(0.097)	(0.213)	(0.233)		(0.077)
$Ret_M Lag$	0.914	5.827^a	-4.913^a	-0.393^a	-0.373^a	−0.103	-0.270^a	-0.287^a
	(1.824)	(1.608)		(0.062)	(0.127)	(0.150)		(0.049)
Control Variables								
TED	0.097^c	-0.233^a	0.330^a	0.013^a	-0.011^b	-0.025^a	0.014^a	0.021^a
	(0.051)	(0.046)		(0.002)	(0.005)	(0.005)		(0.002)
Ban	0.649^a	0.591^a	0.057	0.039^a	0.291^a	0.144^a	0.147^a	0.097^a
	(0.196)	(0.201)		(0.012)	(0.029)	(0.030)		(0.010)
(*Intercept*)	-0.754^a	0.283^a	-1.037^a	-0.119^a	0.087^a	0.180^a	-0.092^a	-0.066^a
	(0.057)	(0.052)		(0.003)	(0.003)	(0.005)		(0.002)
$Adj.\ R^2$	0.736	0.657		0.439	0.648	0.537		0.393

(b) Economic Significance

	EW				*VW*			
	MCI_A	MCI_B	$Diff.$	MCI_{IMB}	MCI_A	MCI_B	$Diff.$	MCI_{IMB}
$Turnover_M$	−0.085	−0.227		0.029	−0.033	−0.056		0.014
VIX	1.145	0.937		0.003	0.056	0.054		−0.008
$UpDown\ Ratio$	0.013	0.030		−0.007	−0.012	−0.015		−0.002
Ret_M	0.093	0.115		−0.003	0.012	0.016		−0.011
$Ret_M Lag$	0.012	0.078		−0.005	−0.005	−0.001		−0.004
TED	0.050	−0.120		0.006	−0.006	−0.013		0.011
Ban	0.046	0.042		0.003	0.021	0.010		0.007

Table 6: *Commonality in liquidity measures*
This table presents summary statistics from time-series regressions of changes in stock liquidity on concurrent, lag and lead market liquidity, as well as a set of control variables estimated for each stock in our sample. Reported are average coefficients together with average t-values for the intercept and three market liquidity coefficients. The section "Sum" reports the average and median of the sum of the three coefficients together with statistics from a t-test and the p-value from a sign-rank test for a difference in mean and median from zero. ΔR^2s indicate how much the inclusion of all three market liquidity variables adds to adjusted R^2. All average t-statistics are reported in parentheses.

	MCI_B	MCI_A	MCI_{IMB}
Intercept	0.093	0.088	−2.189
	(47.309)	(55.716)	(−2.776)
Concurrent	0.531	0.501	0.013
	(60.578)	(76.237)	(0.665)
Lag	0.054	−0.002	0.070
	(6.681)	(−0.400)	(1.076)
Lead	0.036	0.041	0.022
	(3.589)	(7.992)	(0.321)
Sum			
Average	0.620	0.539	0.104
	(31.587)	(46.719)	(0.855)
Median	0.569	0.498	0.001
p-value	0.000	0.000	0.375
Average R^2	0.044	0.049	0.006
Median R^2	0.032	0.043	0.002
Average ΔR^2	0.016	0.015	0.003
Median ΔR^2	0.012	0.013	0.000

Table 7: *Time series of commonality in liquidity measures*

This table presents summary statistics from time-series regressions of changes stock liquidity on concurrent, lag and lead market liquidity, as well as a set of control variables estimated for each stock-semester. Reported are averages and t-values, as well as medians and p-values for the sum of the three market liquidity coefficients. Average ΔR^2s indicate how much the inclusion of all three market liquidity variables adds to adjusted R^2.

	Industrial						**Financial Services**					
	MCI_B			MCI_A			MCI_B			MCI_A		
	Sum		ΔR^2	Sum		ΔR^2	Sum		ΔR^2	Sum		ΔR^2
	Avg	*t*	*Avg*	*Avg*	*t*	*Avg*	*Avg*	*t*	*Avg*	*Avg*	*t*	*Avg*
2002 I	0.86	5.04	0.04	0.60	3.60	0.04	0.97	2.95	0.04	0.49	5.19	0.04
2002 II	0.70	5.16	0.03	1.13	6.60	0.04	0.58	2.29	0.03	0.77	6.55	0.04
2003 I	0.74	7.90	0.03	0.66	5.61	0.04	0.49	2.54	0.03	0.46	4.15	0.03
2003 II	0.02	0.38	0.03	0.21	5.09	0.03	−0.06	−0.64	0.03	0.25	2.33	0.03
2004 I	0.64	7.69	0.03	0.38	6.42	0.04	0.45	3.22	0.03	0.34	3.70	0.03
2004 II	0.52	11.23	0.03	0.19	9.83	0.03	0.75	5.49	0.03	0.49	1.14	0.03
2005 I	0.46	7.27	0.04	0.42	12.41	0.03	0.66	3.58	0.04	0.41	5.52	0.03
2005 II	0.46	10.18	0.03	0.46	8.25	0.03	0.47	4.41	0.03	0.30	6.02	0.03
2006 I	0.18	2.47	0.04	0.17	10.10	0.05	0.31	3.12	0.04	0.42	2.23	0.05
2006 II	0.48	12.18	0.04	0.47	11.46	0.03	0.56	7.86	0.03	0.46	5.43	0.03
2007 I	0.76	26.75	0.12	0.74	36.82	0.13	0.73	6.51	0.11	0.85	12.75	0.11
2007 II	0.62	42.89	0.11	0.33	26.08	0.10	0.64	17.11	0.08	0.41	8.88	0.07
2008 I	0.52	27.36	0.08	0.49	21.24	0.07	0.40	8.70	0.07	0.49	9.59	0.05
2008 II	0.96	43.20	0.22	2.18	24.33	0.11	1.51	18.14	0.20	0.69	14.60	0.08
2009 I	0.70	29.81	0.04	0.82	22.01	0.05	0.74	12.84	0.04	0.76	11.68	0.05
2009 II	1.45	25.82	0.17	1.04	33.04	0.13	1.18	9.86	0.15	1.09	14.80	0.11
2010 I	0.57	20.12	0.05	0.46	21.46	0.04	0.57	7.97	0.05	0.43	7.91	0.05
2010 II	0.12	2.89	0.03	−0.03	0.78	0.03	0.15	3.59	0.02	0.02	−0.62	0.03
2011 I	0.16	7.67	0.04	0.11	6.88	0.03	0.22	4.23	0.04	0.15	2.03	0.03
2011 II	0.32	17.05	0.03	0.29	9.48	0.03	0.25	5.30	0.03	0.23	6.11	0.04
2012 I	0.13	5.79	0.02	0.19	7.69	0.03	0.20	3.67	0.03	0.15	3.11	0.03
2012 II	0.72	21.21	0.08	0.66	17.97	0.06	0.81	9.56	0.08	0.54	7.74	0.06

Table 8: *Predictive pooled regressions.*

This table presents the coefficient estimates from a regression of daily (left panel) and monthly (right panel) percentage stock returns on their own lagged values and the lagged values of MCI_A, MCI_B or their imbalance MCI_{IMB} via pooled OLS. HAC standard errors are reported in parenthesis.

	Daily			*Monthly*		
	MCI_A	MCI_B	MCI_{IMB}	MCI_A	MCI_B	MCI_{IMB}
Intercept	0.0567	0.0573	0.0728	0.8511	0.8301	1.0855
	(0.0019)	(0.0019)	(0.0019)	(0.0360)	(0.0374)	(0.0364)
MCI	0.0070	0.0062	0.0824	0.1290	0.1313	−0.4359
	(0.0007)	(0.0007)	(0.0079)	(0.0139)	(0.0141)	(0.1825)
Lag Return	−0.0170	−0.0174	−0.0167	0.0420	0.0380	0.0386
	(0.0015)	(0.0015)	(0.0015)	(0.0054)	(0.0054)	(0.0055)
Adj. R^2	0.0007	0.0005	0.0003	0.0072	0.0073	0.0016

Table 9: *Summary statistics from the firm-by-firm predictive regressions.*
This table presents some summary statistics for the coefficient estimates of the lagged values of MCI measures from the predictive regressions of daily (upper panel) and monthly (lower panel) stock returns on their own lagged values and each of these variables, separately. The regressions are estimated via least squares separately for each stock with at least 63 observations for daily and 24 for monthly analysis. 5^{th}, 50^{th} and 95^{th} denote the corresponding quantiles of the distribution of coefficient estimates. $pos(\&sig)$ and $neg(\&sig)$ denote the percentage of coefficients that are positive and negative (and significantly so), respectively.

(a) Prediction of daily returns

	$Mean$	$StdDev$	5^{th}	50^{th}	95^{th}	pos	$pos\&sig$	neg	$neg\&sig$
MCI_A	0.29	1.22	−0.14	0.06	1.49	0.80	0.25	0.20	0.01
MCI_B	0.22	0.99	−0.19	0.02	1.32	0.68	0.14	0.32	0.03
MCI_{IMB}	0.07	0.76	−0.89	0.09	0.95	0.62	0.12	0.38	0.03

(b) Prediction of monthly returns

	$Mean$	$StdDev$	5^{th}	50^{th}	95^{th}	pos	$pos\&sig$	neg	$neg\&sig$
MCI_A	2.55	23.40	−7.75	0.69	16.57	0.76	0.29	0.24	0.03
MCI_B	2.40	20.92	−6.32	0.35	16.15	0.69	0.23	0.31	0.05
MCI_{IMB}	−1.40	18.67	−27.51	0.28	16.18	0.51	0.07	0.49	0.08

Table 10: *Daily portfolio sorts.*

This table presents mean returns (Panel (a)) and four-factor alphas (Panel (b)) of portfolios formed on a daily basis by sorting based on different *MCI* measures. At the end of each trading day, stocks are assigned to a (equally weighted (EW) or value-weighted (VW)) portfolio according to the quintile of their *MCI* measure and held until the end of the next trading day, when new portfolios are defined. Portfolio 1 consists of stocks with the lowest *MCI* measures, and Portfolio 5 consists of those with the highest *MCI* measures. 5-1 is the portfolio formed by having a long position in Portfolio 5 and a short position in Portfolio 1. Four-factor alphas are computed based on the four-factor model of Carhart (1997). All numbers are in basis points. *a*, *b* and *c* denote significance at the 1%, 5% and 10% significance levels, respectively, which are omitted for returns on Portfolios 1 to 5 in Panel (a) because they are all significant.

(a) Returns

	Bid		*Ask*		*Imb*	
Portfolio	*EW*	*VW*	*EW*	*VW*	*EW*	*VW*
1	4.310	3.140	3.710	3.010	3.170	3.840
2	5.730	5.130	4.580	4.900	4.400	3.330
3	6.610	6.150	6.610	6.900	6.310	4.140
4	7.240	6.250	6.560	7.020	7.380	3.370
5	7.630	4.970	10.060	7.990	10.260	3.910
5-1	3.320	1.830	6.360^{a}	4.980^{b}	7.090^{a}	0.070

(b) Four-factor Alphas

	Bid		*Ask*		*Imb*	
Portfolio	*EW*	*VW*	*EW*	*VW*	*EW*	*VW*
1	0.957^{c}	0.084	0.308	−0.058	-1.729^{b}	1.079
2	2.288^{a}	1.836^{b}	1.120	1.667^{b}	−0.251	−0.191
3	2.242^{a}	2.129^{a}	2.062^{a}	2.801^{a}	2.333^{a}	1.236^{c}
4	2.423^{a}	1.748^{b}	1.726^{b}	2.630^{a}	3.238^{a}	0.051
5	1.630^{c}	−0.558	4.325^{a}	2.615^{a}	5.948^{a}	0.039
5-1	0.673	−0.642	4.017^{a}	2.673^{a}	7.677^{a}	−1.040

Table 11: *Monthly portfolio sorts.*
This table presents mean returns (Panel (a)) and four-factor alphas (Panel (b)) of portfolios formed on a monthly basis by sorting based on different *MCI* measures. At the end of month M (depending on the average *MCI* value in M), stocks are assigned to a (equally weighted (EW) or value-weighted (VW)) portfolio according to the quintile of their *MCI* measure and held until the end of month M+1, when new portfolios are defined. Portfolio 1 consists of stocks with the lowest *MCI* measures and Portfolio 5 consists of those with the highest *MCI* measures. 5-1 is the portfolio formed by having a long position in Portfolio 5 and a short position in Portfolio 1. Four-factor alphas are computed based on the four-factor model of Carhart (1997). All numbers are in percentages. *a*, *b* and *c* denote significance at the 1%, 5% and 10% significance levels, respectively.

(a) Returns

	Bid		*Ask*		*Imb*	
Portfolio	*EW*	*VW*	*EW*	*VW*	*EW*	*VW*
1	0.693	0.489	0.668	0.490	1.100	0.538
2	0.898	0.853	0.847	0.828	0.898	0.555
3	1.110	1.076	1.107	1.127	0.929	0.740
4	1.124	1.070	1.219	1.113	1.144	0.644
5	1.486	1.182	1.472	1.233	1.238	0.804
5-1	0.792	0.693	0.805	0.743	0.138	0.267

(b) Four-factor Alphas

	Bid		*Ask*		*Imb*	
Portfolio	*EW*	*VW*	*EW*	*VW*	*EW*	*VW*
1	0.173^{c}	0.094	0.140	0.094	0.278^{c}	-0.015
2	0.197	0.191	0.161	0.182	0.136	0.061
3	0.301^{b}	0.326^{b}	0.306^{b}	0.382^{a}	0.149	0.280^{b}
4	0.181	0.200	0.263	0.226	0.338^{a}	0.185
5	0.382^{b}	0.198	0.362^{c}	0.217	0.325^{c}	0.244
5-1	0.209	0.104	0.222	0.123	0.047	0.259