Weather and Time Series Determinants of Liquidity in a Limit Order Market*

Juhani Linnainmaa[†] Ioanid Roşu[‡]

October 13, 2009

Abstract

When liquidity is measured by the bid-ask spread or price impact, markets with more trading activity are typically more liquid than markets with less trading activity. But showing a causal connection from trading activity to spreads is difficult because these variables are endogenous. In the case of Finland's fully electronic limit order market, we use deseasonalized sunshine as an instrument for trading activity, and find that indeed higher trading activity causes lower spreads in the time series. We introduce another instrument for spreads and show that causality runs the other way as well: lower bid-ask spreads invite more trading activity. By using the lagged CBOE Volatility Index as an instrument, we also find that an exogenous increase in intra-day volatility causes larger spreads.

JEL Classification: G1

Keywords: Market microstructure, price impact, instrumental variables.

*We thank Alan Bester, Doug Diamond, John Heaton, Chris Hansen, Atif Mian, and Duane Seppi for helpful comments and suggestions. We are also grateful to participants at the AFA 2009 Meeting, the Conference on Trading Frictions in Asset Markets, UC Santa Barbara; and to seminar audiences at Chicago Booth and the Academy of Economic Studies in Bucharest.

[†]University of Chicago, Booth School of Business, Email: juhani.linnainmaa@ChicagoBooth.edu.

[‡]University of Chicago, Booth School of Business, Email: ioanid.rosu@ChicagoBooth.edu.

1 Introduction

Recent studies suggest that liquidity, as measured by bid-ask spreads, price impact, or price reversals, is an important determinant of expected returns.[1] Because there exists significant cross sectional and time series variation in liquidity, it is important for asset pricing research to understand the sources of this variation.

There is a large literature on the cross sectional determinants of liquidity. For example, Stoll (2000) reports that about 80% of the cross sectional variation in bid-ask spreads is due to easily measurable stock characteristics.[2] Spreads are lower for stocks with greater trading volume, lower return volatility, higher price, and smaller trading imbalances. The question is how to interpret these results. For example, do the results imply that a stock has a lower bid-ask spread *because* it has more trading volume? Or does it have more trading volume because lower spreads invite more trading activity?

An OLS regression of bid-ask spreads on trading activity, or vice versa, fails to identify the causal link between them. The resulting problem is akin to the lack of identification in a regression of prices on quantities. By using instrumental variables for both trading activity and bid-ask spreads, we show that the causal relation works in both directions. This paper is, to the best of our knowledge, the first to tackle this endogeneity problem.

If we believe the relations we investigate are causal, then they should also hold in the time series. But despite significant daily variation in the liquidity of stocks, the time series determinants of liquidity have been less studied in the literature.[3] Does the liquidity change because of asymmetric information or because of liquidity shocks? Is this variation market-wide or stock-specific? Just as in the cross section, it is difficult to find convincing answers to these questions. First, many potential determinants of liquidity, such as the extent of asymmetric information, are not observable. And second, time series studies of liquidity are bound to suffer from endogeneity problems as well.[4]

---

[1] See, for example, Pástor and Stambaugh (2003), Acharya and Pedersen (2005), and the references therein.

[2] See also Stoll (2003).

[3] Chordia, Roll, and Subrahmanyam (2000) study correlated movements in liquidity, and Chordia, Roll, and Subrahmanyam (2001) analyze patterns in the time series movements of aggregate stock liquidity.

[4] Some papers try to get around the endogeneity problem by considering a structural model of trades and

This paper studies the time series determinants of liquidity by using detailed data from the Helsinki Stock Exchange (HEX), Finland's fully electronic limit order market.[5] We use three different measures of liquidity: the percentage bid-ask spread, the effective spread, and the price impact of a trade. In order to take advantage of all the information in the limit order book, we modify the last two of these measures. Instead of estimating the price impact of a trade by running an endogenous regression of prices on order flow, or by measuring the effective spread from realized trades, we read these measures directly from the order book.[6]

We first find that trading activity is an important *cause* of liquidity. This is predicted by recent theoretical work on limit order markets, although the intuition is already present in Demsetz (1968).[7] The intuition for order driven markets is that a higher trading activity lowers the costs of waiting for limit order traders who compete against each other. The decrease in waiting costs makes traders more willing to place their limit orders closer to each other. This change in liquidity provision behavior in turn lowers both the bid-ask spread and price impact.

We argue that (deseasonalized) sunshine is a good instrument for trading activity. First, we find that household arrival rates are affected considerably by changes in the weather. This can be explained rationally by the opportunity cost of trading: if the weather is good, those individuals who are not time-constrained to trade may rationally postpone trading. Second, it is difficult to believe that investors are affected by the weather *after* they decide to participate in the market. Consistent with this intuition, we find that weather does not affect the order choice of institutions, who presumably have less latitude about market participation. Third, because we study non-agriculture stocks, the variation in weather is uncorrelated with both

quotes. See, for example, Glosten and Milgrom (1985), George, Kaul, and Nimalendran (1991), Huang and Stoll (1997), Bollen, Smith, and Whaley (2004). These papers typically decompose the spread into several components, including asymmetric information, inventory-holding costs, and order-processing costs. They find that find that the component due to asymmetric information is small (less than 15%), while the "order processing" or "inventory-holding" costs are the largest components.

[5] A limit order market, or an order driven market, is a market where trading takes place according to a limit order book. Nowadays most markets around the world are at least partially order driven.

[6] Our price impact measure estimates the average *instantaneous* price impact that a hypothetical market order has on the limit order book. The difficulty comes from the fact the instantaneous price impact can be infinite for large enough market orders: there is no depth beyond a certain level. Our solution is to average over the *inverse* price impact, which even if it becomes zero it poses no problem in the averaging process.

[7] See, for example, Foucault, Kadan, and Kandel (2005) and Roşu (2009a).

the discount rate and cash flow news of our sample stocks.

By using weather as an instrument, we find that higher trading activity causes lower spreads. The effect is both statistically and economically significant. In the first stage of the instrumental variables (IV) regression, a one standard deviation shock to the weather variable increases limit order arrival rates by 1.5%, market order arrival rates by 2.3%, and all trading activity by 1.7%.[8] We show that these effects are largely due to households' sensitivity to weather.

In the second stage of the IV regression we find that a quarter standard deviation *exogenous* shock to (the natural logarithm of) trading activity decreases percentage bid-ask spreads by 12.35%. This estimate is more than three times as large as the estimate obtained from an OLS regression. This effect is economically significant. For example, the median percentage bid-ask spread across our sample stocks is 2.27%. If a stock trades at EUR 30, this corresponds to a bid-ask spread of EUR 0.68. Then, a quarter standard deviation exogenous shock to trading activity would cause the bid-ask spread to narrow to EUR 0.60.

One should not infer from these results that causality runs only from trading activity to spreads. It may be that smaller spreads also generate higher trading activity. In their empirical study of the Paris Bourse, Biais, Hillion, and Spatt (1995) find that smaller bid-ask spreads attract more market orders and fewer limit orders. Thus, on balance, it is not clear which way spreads influence trading activity. A simple OLS regression in our case shows a negative correlation between spreads and trading activity. We study causality by constructing another instrument for the changes in bid-ask spread. We first count the number of cases where a limit order trader at the bid or ask changes his order to a market order. This affects the spread mechanically, but it should not affect the trading activity directly. We then define the instrument as the number of limit-to-market order modifications at the bid or ask, normalized by the total number of such modifications. This normalization guards against the possibility that trading activity uniformly increases the number of limit-to-market order modifications. An IV regression shows that indeed smaller spreads cause larger trading activity: e.g., if the

[8]We find similar results if we study the variation of weather within the same day. For each firm and on each day we define the company-specific weather by measuring weather closest to the weather station where the most shareholders of that company reside. Then firms for which the company-specific weather is better have significantly lower trading activity.

bid-ask spread of a EUR 30-stock were to decrease exogenously from EUR 0.68 to EUR 0.60, trading activity would increase by approximately 28%.

Once we establish the validity of weather as an instrument for trading activity, we turn the analysis to other determinants of market liquidity. Several market microstructure models suggest that a key determinant of liquidity is the volatility of the fundamental value of the asset.[9] Some empirical studies, such as Ranaldo (2004), show that high price volatility is correlated with high spreads. But this is an endogenous regression, and it uses price volatility rather than the (unobserved) fundamental volatility. Therefore, the result may reflect reverse causality: price volatility is relatively higher in markets with higher bid-ask spreads.

In order to study the causality from volatility to spreads, we use the lagged U.S. CBOE Volatility Index (VIX) to instrument for price volatility. VIX is almost surely exogenous to the Finnish stock market: trading in Finland does not significantly affect VIX. Moreover, one could plausibly argue that VIX affects liquidity in Finland only via its effect on fundamental volatility and trading activity. Our IV regression results are both statistically and economically significant: a quarter standard deviation shock to volatility increases the percentage bid-ask spread by 14.1%.

In addition to the bid-ask spread, we study the effect of trading activity and volatility on other liquidity measures constructed from the limit order book: the effective spread, and an instantaneous price impact measure. The results for effective spreads are even stronger than those for the bid-ask spread, indicating that the effective spread—which ignores very small orders placed at the bid or ask—is a less noisy measure of the true bid-ask spread.

We find that higher trading activity also lowers the price impact. A quarter standard deviation exogenous increase in log-trading activity decreases our price impact measure by 0.045, which represents about a quarter standard deviation change in the price impact. If we consider a stock with the median price impact in the sample, $\lambda = 2.063$, this corresponds to an instantaneous price impact of 8.03% per 1,000 shares. If price impact decreases by 0.045 to 2.018, this corresponds to an average price impact of 5.68% per 1000 shares. (These figures are high because the instantaneous price impact measures the price impact of a *hypothetical*

[9] See Kyle (1985), Glosten and Milgrom (1985), Foucault (1999), and Roşu (2009b).

trade. In reality, traders rarely submit large market orders when there is insufficient liquidity in the book. Instead, they would submit a small order and wait for the limit order book to be replenished before submitting another small order.)

We find that the results on the price impact measure are generally stronger than the results on the bid-ask spread and the effective spread. This result is intuitively appealing: because it takes into account all the existing depth in the order book, the price impact of a trade is a better proxy for liquidity than the bid-ask spread and the effective spread, which only measure liquidity around the bid and ask. We often use bid-ask spreads not because they best proxy for what we understand as "liquidity" but because they are easily available.

The paper is organized as follows. Section 2 discusses the theory on the determinants of liquidity in limit order markets. Section 3 introduces the data and defines the liquidity measures. Section 4 discusses our empirical strategy and the instrumental variables. Section 5 presents the results. Section 6 concludes.

2 Determinants of Liquidity in Order Driven Markets

A large body of empirical work suggests that liquidity is important for asset pricing. An asset is said to possess *liquidity* if it can be traded quickly, in large quantities, and without significantly affecting the market price. Black (1971) and Kyle (1985) define a liquid market as being: (1) tight, if bid-ask spreads are small; (2) deep, if the price impact of a transaction is small; and (3) resilient, if prices recover quickly after large orders move the price away from fundamentals. We group (1) and (2) under the name of "market depth." In this paper, we study liquidity along the market depth dimension by analyzing the bid-ask spread, the effective spread, and the price impact of a transaction. The market depth component of liquidity is known to be a significant determinants of the cross section of asset returns.[10]

To discuss the determinants of liquidity, we start with Kyle (1985). He proposes an asymmetric information model with a monopolist insider, a competitive liquidity provider (market

[10] See, for example, Brennan and Subrahmanyam (1996) and Acharya and Pedersen (2005). The price reversal liquidity measure of Pástor and Stambaugh (2003) suggests that market resilience is also important for asset prices.

maker), and noise traders. In this model the noise trading activity and the fundamental volatility are the two determinants of liquidity. The model predicts that an increase in noise trading activity decreases the price impact of a trade because it makes the order flow less informative: the informed trader can better hide his private information. By contrast, an increase in the fundamental volatility increases the price impact of a trade because it makes the order flow more informative: the market maker has less prior knowledge about the true value.[11]

Kyle's arguments depend crucially on the assumptions that the insider has a monopoly over private information and that the market maker is competitive. With more than one insider it is not clear whether Kyle's results still hold. If more insiders arrive when there is more noise trading—as Admati and Pfleiderer (1988) predict—a higher trading activity may actually lead to a higher price impact. Also, as shown in the empirical study of Sandås (2001), the second assumption of perfectly competitive markets fails in limit order markets. This suggests that the amount of private information, as measured by the ratio of informed investors to uninformed investors, is also an important determinant of liquidity. We call this ratio the *information ratio*. This gives us three determinants of liquidity derived from asymmetric information: noise trading activity, fundamental volatility, and the information ratio.

Theoretical models of limit order markets show that all three potential determinants of liquidity are important. First, in order driven markets a higher noise trading activity usually leads to smaller spreads. This is true in asymmetric information models such as Glosten (1994) as long as the arrival rate of limit order traders does not depend on the amount of noise trading. But there is another explanation that instead of asymmetric information it relies on trading costs, and in particular on waiting costs.[12] First, suppose that some traders are impatient and demand liquidity by using market orders and that some are patient and supply liquidity by using limit orders. Then, the spread between limit orders determines the

[11]In a dealer market setup, Glosten and Milgrom (1985) propose a similar explanation of the effect of fundamental volatility on the bid-ask spread. When the uncertainty regarding the true value of the asset is high, so is the difference between the expected value conditional on someone buying and the expected value conditional on someone selling. This is precisely the bid-ask spread, which therefore should also be high.

[12]See Foucault, Kadan, and Kandel (2005) and Roşu (2009a), but also Goettler, Parlour, and Rajan (2005), Roşu (2009b), and the earlier papers of Demsetz (1968), Cohen, Maier, Schwartz, and Whitcomb (1981), Handa and Schwartz (1996), Parlour (1998), Seppi (1997), and Foucault (1999).

compensation of the patient traders. Therefore, when trading is more frequent, limit order traders wait less and thus accept to be compensated with smaller spreads.[13] Thus, an increase in trading activity (as measured by the sum of all arrival rates) decreases bid-ask spreads and increases market liquidity.

Second, according to Foucault (1999) and Roşu (2009b) fundamental volatility also affects liquidity. Foucault (1999) proposes a model of limit order markets that focuses on the "picking off" risk that stale limit orders face as soon as new public information about the asset arrives. In this model, a high fundamental volatility leads to a higher probability of limit orders being picked off. Then the liquidity providers submit less aggressive limit orders, which widens the bid-ask spread. In Roşu (2009b) bid-ask spreads are a measure of the public uncertainty about the fundamental value, so an increase in fundamental volatility widens the bid-ask spread and increases the price impact of a trade.

Third, Roşu (2009b) shows that the information ratio is determinant of liquidity and that a higher information ratio leads to smaller spreads. This is in contrast with Glosten (1994), in which finds a higher information ratio leads to higher spreads. The difference comes from the fact that the former is a dynamic model in which the fundamental value changes over time and informed traders can choose between limit and market orders. (In Glosten (1994) informed traders always submit market orders.) The public price converges faster to the fundamental value in the presence of more informed traders. This lowers the uncertainty about the fundamental value of the asset and thus decreases spreads. Moreover, a higher information ratio makes prices revert more quickly to fundamentals, making the market more resilient.

Because the focus of this paper is on market depth (spreads and price impact) and not on market resiliency, in our empirical study we consider only the first two variables: noise trading activity and fundamental volatility.

[13]Here, the word "spread" refers to the distance between two orders on the same side of the book and not to the bid-ask spread.

3 Data and Liquidity Measures

3.1 Helsinki Stock Exchange

Trading on the Helsinki Stock Exchange (HEX) starts each day with an opening call at 10:10 a.m. Orders that are not executed at the opening remain in the consolidated limit order book and form the basis for the continuous trading session. This trading session takes place between 10:30 a.m. and 5:30 p.m. in a fully automated limit order book. After-hours trading (5:30–5:45 p.m.) takes place after the continuous trading session and again the next morning (9:30–10:00 a.m.). Brokers can only report pre-negotiated trades in the after-hours session.

The minimum tick size is EUR 0.01. The HEX trading system displays the five best price levels of the limit order book on both sides. The public can view this book in a market-by-price form while financial institutions also receive market-by-order feed. A market-by-price book displays the amount of shares outstanding at the five best price levels on both sides of the market. A market-by-order book shows each order separately and also shows which broker submitted each order. HEX has no designated market makers or specialists. Investors trade by submitting limit orders. An investor who wants immediate execution must place the limit order at the best price level on the opposite side of the book. We call these marketable limit orders "market orders" in this study. An investor who wants to buy or sell more shares than what is currently outstanding at the best price level must submit separate orders for each price level. If a limit order executes against a smaller order, the unfilled portion stays on the book as a new order. Time and price priority between limit orders is enforced. For example, if an investor submits a buy order at a price level that already has other buy orders outstanding, all the old orders must execute before the new order gains priority.

3.2 Limit Order Data

We use the limit order data from the Helsinki Stock Exchange in this study. These data are derived from the supervisory files from the HEX from July 10, 2000 through October 23, 2001. Since the data come directly from the actual trading platform, they are highly reliable.

Each entry in the limit order data is a single order entered into the system; it contains a unique order identifier, date- and time-stamps, a session code, a code for the brokerage firm submitting the order, a trade type indicator (i.e., whether the trade is a regular trade, a (pre-negotiated) block trade, or an odd-lot trade), and the price and volume of the order.

All entries also contain a set of codes for tracking the life of an order. An order can expire, be partly or completely filled, or modified. We use these data to reconstruct the limit order book for each second of every trading day for all stocks.

The limit order data do not have identities of the traders who submit the orders. However, another data set, the Finnish Central Securities Depository registry, contains complete trading records in all publicly traded Finnish stocks. (Grinblatt and Keloharju (2000) provide more details about this data set.) This registry can be combined with the limit order data to obtain identities for those orders that execute at least partially and thus appear also in the FCSD registry. Although the registry has no information about orders that do not execute, Linnainmaa (2009) shows that order characteristics can be used to calculate very accurate approximate trader identities also for these orders. Linnainmaa estimates a probit model to approximate identities for unfilled orders, using order characteristics such as order size, initial order placement, and broker identities as explanatory variables. The model gets the identity of the investor right 85.4% of the time for executed orders. We use the same method this study to classify orders to those originating from household investors and to those originating from institutional investors.

The full sample contains 166 companies, some of which have dual share classes, giving us a total of 185 stocks. We limit the sample to 88 companies by requiring that each company has at least an average of five trades (marketable limit orders) per day.

3.3 Liquidity Measures

The concept of liquidity goes well beyond the size of the bid-ask spread. Black (1971, p. 29–30) defines an asset as liquid if "it can be sold in a short time, at a price not too much below the price the seller would get if he took plenty of time to sell the asset." Most research on liquidity is hindered by data restrictions. For example, the CRSP data have information only on the closing bid-ask spread. A switch to intraday transaction data sets allows one to construct more intricate liquidity measures. However, another problem that arises is market fragmentation: it is often difficult to characterize liquidity because there is no centralized exchange. The advantage of studying a market with a consolidated limit order book is that we have perfect information about *instantaneous* liquidity: "if an investor were to submit a buy market order for x shares, the average purchase price would be y." Although no data can fully reveal the amount of latent liquidity—i.e., limit orders that have not been submitted because some traders are unwilling to disclose their intentions—the rich data from a limit order market make it possible to consider different facets of liquidity.

We define the following variables:

- (Percentage) bid-ask spread: The log-difference between the best ask price and the best bid price in the limit order book. We compute this bid-ask spread every 10 seconds of the trading day and use the average as the measure of the daily bid-ask spread.

- (Percentage) effective spread: In dealer or specialist markets this is usually defined as twice the difference between the trade price and the time-of-trade bid-ask spread midpoint.[14] In a pure order driven market, trades are generated by market orders, so we could define the effective spread as the execution price of the market order. The problem with this measure is that the effective spread is not defined on days when there are no trades, thus making the comparison with the other liquidity measures more difficult. Instead, we define the effective spread as the price impact of a market order of average size.

  More precisely, we start by computing for each stock the average size of a trade over the sample period. Then, at each time we define the buy trade price as the average price at which a hypothetical marketable buy order of the average size would execute. For example, suppose that at the ask side of the book there are 100 shares for EUR 10 and

[14] When trades take place within the spread because the dealers or specialists arrange price improvement, the effective spread is smaller than the bid-ask spread. When large orders fill at prices outside the bid-ask spread, the effective spread is greater than the bid-ask spread.

200 shares for EUR 11. If the typical market order size in this stock is 150 shares, 100 shares of this market order would execute at 10 and the rest (50 shares) would execute at 11. Hence, the buy trade price would be $(100*10+50*11)/(100+50) =$ EUR 10.33. We compute the sell trade price in a similar manner. If both these prices exist (i.e., there is enough liquidity in the book to accommodate a regular-sized market order), we record the effective spread as the buy trade price minus the sell trade price. Finally, we define our daily effective spread measure to be the average effective spread on that day, sampled every 10 seconds.

- Price impact: We define the price impact by studying how much the price would change if the investor placed a market order that would clear some existing limit orders in the book. Suppose there are K orders on the ask side (the measure is symmetric on the bid side). We truncate our definition of price impact by looking only at the first 10 levels in the book. For each $i = 1, 2, \ldots$ denote by p_i the price corresponding to the i'th level counting from the ask price. Denote by d_i the market depth at that price in dollars, i.e., the dollar value of all sell limit orders at the corresponding price.

  Now suppose a trader submits a market order of size $d_1 + \cdots + d_k$. Then the ask price moves from p_1 to p_{k+1}. The Kyle lambda measure of price impact would then be

$$\lambda_k = \frac{p_{k+1} - p_1}{d_1 + \cdots + d_k} = \frac{(p_2 - p_1) + \cdots + (p_{k+1} - p_k)}{d_1 + \cdots + d_k}.$$

  One could define an average price impact measure by taking the average of all λ_k for $k = 1, \ldots, K$. But there are illiquid stocks with few offers in the book (K small), and for those the price impact for market orders larger than $d_1 + \cdots + d_K$ would be undefined. The solution is to consider the inverse $\frac{1}{\lambda_k} = \frac{d_1+\cdots+d_k}{(p_2-p_1)+\cdots+(p_{k+1}-p_k)}$, which is usually called "market depth." This can be seen to be a weighted average of $\Delta_i = \frac{d_i}{p_{i+1}-p_i}$. The depth measure Δ_i works well because it equals zero when there is no depth at a certain level. We can approximate $\Delta_i \approx \frac{d_i/p_i}{\log(p_{i+1}/p_i)} = \frac{s_i}{\log(p_{i+1}/p_i)}$, where s_i is the total number of shares offered at p_i. The problem with this last definition is that s_i is a large number, and $\log(p_{i+1}/p_i)$ can be a very small number, both of which may generate large values of Δ_i. So instead of Δ_i we consider $\ln(1 + \Delta_i)$. Finally, we need to define p_{K+1} if

$K < 10$. Normally p_{K+1} would equal infinity, since there is no depth there, but we set $p_{K+1} = 2p_K$.

We define the price impact in the following way. Let K again be the number of offers in the book, or 10, whichever is smaller. Define

$$\Delta_i \;=\; \begin{cases} \ln\left(1 + \frac{s_i}{\log(p_{i+1}/p_i)}\right) & \text{if } i = 1, \ldots, K. \\ 0 & \text{if } i = K+1, \ldots, 9 \end{cases}$$

Define the market depth on the ask side by $\Delta_{\text{ask}} = \frac{\Delta_1 + \Delta_2 + \cdots + \Delta_9}{9}$. The market depth on the bid side Δ_{bid} is defined similarly, but one must replace $\ln(p_{i+1}/p_i)$ by $\ln(p_i/p_{i+1})$. Then market depth is defined as $\Delta = (\Delta_{\text{ask}} + \Delta_{\text{bid}})/2$, and the price impact measure is defined by

$$\lambda \;=\; \ln\big(1 + 16.3 - \Delta\big).$$

Here we used the maximum value $\Delta_{\max} = 16.3$ that Δ takes across the whole panel data set. (This maximum is attained by Nokia, the most actively traded stock in the sample.) We compute this price impact measure for every second of the trading day and use the average as the measure of the daily price impact.

We do not define λ by using the direct inverse: $\lambda = \frac{1}{\Delta}$ for two reasons. First, on days when there is no trading, $\Delta = 0$, so the price impact in that case would be infinite. Second, this latter measure of price impact ignores the variation in the price impact measure when companies are not too illiquid. To see this, consider the median company in this sample (Satama Interactive), which has an average market depth $\bar{\Delta} \approx 9.6$. The most liquid company is Nokia, with $\bar{\Delta} \approx 15.1$. Then the difference between $\frac{1}{9.6} \approx 0.10$ and $\frac{1}{15.1} \approx 0.07$ is very small compared with the difference in $\frac{1}{\Delta}$ for two illiquid companies (which have Δ close to zero).

Table 1 reports summary statistics for various liquidity measures. To interpret the figures from this table, recall that the price impact measure λ_t on a given day t is defined as $\lambda_t = \ln(17.3 - \Delta_t)$, where $\Delta_t = (\Delta_{\text{ask},t} + \Delta_{\text{bid},t})/2$ is the average market depth for that day. Then the price impact λ for a stock is the average λ_t. In Table 1 we see that the median price impact is 2.063, which refers to a stock called Satama

Interactive. This corresponds to an implied market depth of $\Delta = 17.3 - \mathrm{e}^{\lambda} = 9.43$, although by Jensen's inequality this is different from the average of all Δ_t. For Satama Interactive, we see that in fact the average market depth is 9.621.

First, for simplicity, assume that all Δ_i from the definition of Δ are equal, i.e., the price impact function is linear. Then $\Delta_i = \ln(1 + \frac{s_i}{\log(p_{i+1}/p_i)}) = 9.621$. From this we compute $\frac{\log(p_{i+1}/p_i)}{s_i} = 1/(\mathrm{e}^{9.621} - 1) \approx 6.63 \times 10^{-5}$. Since s_i is the number of shares on level i, it follows that 1000 shares traded in this stock make the price change instantaneously on average by $0.0663 = 6.63\%$. This means that the median stock in Finland is quite illiquid.

But the computation is deceptive, because the stock does not have a positive number of limit orders on all the nine levels used to compute λ. In fact, on average, Satama Interactive has only 7.5 price levels in each side of the limit order book. This means that the average *non-zero* Δ_i equals $\frac{9}{7.5} \times 9.621 = 11.545$. This corresponds to a price impact per share of $\frac{\log(p_{i+1}/p_i)}{s_i} = 1/(\mathrm{e}^{11.545} - 1) \approx 9.68 \times 10^{-6}$. This means that a market order for 1000 shares that does not consume more than the 7 or 8 levels in the limit order book for Satama Interactive would have a price impact of approximately $0.00968 \approx 1\%$. However, an order that consumes more than 8 levels in the book would typically have an infinite price impact. Notice then that our price impact measure is constructed in such a way that it can average out even over infinite values (essentially the measure is a harmonic average, which has this property).

To simplify the discussion, we use the implied market depth to estimate the per-share price impact. For example, if the price impact measure is $\lambda = 2.063$, then the implied market depth equals $\Delta = 17.3 - \mathrm{e}^{\lambda} = 9.43$, and this corresponds to a per-share price impact equal to $\frac{\log(p_{i+1}/p_i)}{s_i} = 1/(\mathrm{e}^{9.43} - 1) \approx 8.03 \times 10^{-5}$. This means, with all the caveats discussed above, that 1000 shares traded in this stock lead on average to a price impact of 8.03%. If, for example, the price impact were to decrease to $\lambda = 2.018$, this would correspond to a price impact of about 5.68% per-1000 shares. Indeed, the implied market depth equals $\Delta = 17.3 - \mathrm{e}^{\lambda} = 9.78$, which corresponds to a per-share price impact equal to $1/(\mathrm{e}^{9.78} - 1) \approx 5.68 \times 10^{-5}$.

- Trading activity: For each day, denote by BMO the number of buy market orders, BLO the number of buy limit orders, by SMO the number of sell market orders, and by SLO the number of sell limit orders. Then trading activity on that day equals $BMO + BLO + SMO + SLO$. In our empirical study we use the natural logarithm of trading activity (plus one), since its distribution is closer to being normal.

- Trading competition: Define first competition on the buy side: Buy Competition = $(1 + BLO)/(1 + SMO)$. Similarly, competition on the sell side is: Sell Competition = $(1 + SLO)/(1 + BMO)$. Then Competition is the average of Buy and Sell Competition. This variable measures competition in liquidity supply.

- Midpoint price: This is the daily average of the midpoint between the ask price and the bid price.

- Midpoint volatility: This is the standard deviation of the natural logarithm of the intraday midpoint price.

- Order autocorrelation: This measures the diagonal effect of Biais, Hillion, and Spatt (1995): after a certain type of order, an order of the same type is more likely. If t indicates the transaction number, then define theoretical order autocorrelation by $\big(P(BMO_{t+1}|BMO_t) + P(BLO_{t+1}|BLO_t) + P(SMO_{t+1}|SMO_t) + P(SLO_{t+1}|SLO_t)\big)/4$. If order flow was not autocorrelated, then this would equal $\big(P(BMO_{t+1}) + P(BLO_{t+1}) + P(SMO_{t+1}) + P(SLO_{t+1})\big)/4 = 0.25$. If the order flow was autocorrelated, we should expect a number higher than 0.25. In practice, we estimate these autocorrelations in the following way. First, count all the pairs of successive order occurrences (BMO_t, O_{t+1}), where O_{t+1} can be any order (BMO,BLO,SMO,SLO). Then count the number of occurrences (BMO_t, BMO_{t+1}). Divide this by the total count, and this is our proxy for $P(BMO_{t+1}|BMO_t)$. The other conditional probabilities are estimated in the same way. If we lack data to estimate one of these conditional probabilities, we use the average of the three remaining conditional probabilities to measure autocorrelation. For example, if there are no data on $P(BMO_{t+1}|BMO_t)$, we compute the order autocorrelation as $\big(P(BLO_{t+1}|BLO_t) + P(SMO_{t+1}|SMO_t) + P(SLO_{t+1}|SLO_t)\big)/3$.

Table 1 shows that the mean order autocorrelation in our sample is 0.3, which is larger than 0.25. This confirms that the diagonal effect is present in the Finnish stock market as well.

- Order imbalance: This measures the imbalance between the number of buy market orders (BMO) and sell market orders SMO. Define the daily imbalance by $(BMO - SMO)/(BMO + SMO)$ if there is at least a buy or sell market order for that day, and zero otherwise.

<< TABLE 1 ABOUT HERE >>

Table 1 reports summary statistics for these variables. There is considerable cross sectional variation in these measures across our sample stocks. The average bid-ask spread is less than 0.72% for 5% of the stocks traded on the HEX but above 5.87% for the 5% of the stocks at the other end of the spectrum. Nevertheless, differences in stock-specific means does not explain all the variation observed in the data. The last column shows that such cross sectional variation in means accounts for about 49% of the variation in the panel data. The remaining variation arises from time series changes. The adjusted standard deviation reported in the second to the last column reports the amount of residual variation after having removed stock-specific averages. The R^2 measure reported in the last column is obtained from regressing each measure against all firm fixed effects. This residual variation is the relevant measure of standard deviation for our time series study.

4 Empirical Strategy: Instrumental Variables

One of the main purposes of this paper is to analyze the effect of trading activity on liquidity. But both market liquidity and trading activity are endogenous variables: more traders arrive to a liquid market, and an increase in the number of traders improves liquidity. We need to use instrumental variables to identify causal effects.

The first instrument variable we propose is weather, as measured by the deseasonalized number of hours of sunshine per day. We analyze households and institutions separately to study the effect of weather on order submission choices. We suggest that weather should be correlated with the behavior of households, but not with the behavior of institutions: If a household is not constrained to trade immediately, an exogenous shock in weather may cause it to postpone trading. By contrast, institutional traders have less discretion over whether to show for work or not. This gives us an opportunity to test the validity of weather as an instrument: if weather affected institutions at all, then this would suggest that weather might be correlated with some omitted variable. This would pose a problem in the second stage, because we could not be certain how weather affects liquidity: are spreads smaller because weather kept some people from arriving to the market, or are they smaller because weather affects the behavior of the people who are in the market, via some unknown channel?

A priori, we expect weather to be a good instrument for trading activity. First, weather should be correlated with trading activity: we expect some individual investors who value leisure to rationally take advantage of good weather and not trade as much. Second, it is plausible that weather does not affect the behavior of investors once they do arrive to the market.[15] Also, none of the stocks traded on the Helsinki Stock Exchange are agricultural. Thus, the variation in day-to-day weather should not be correlated with changes in either expected future cash flows or discount rates. Because of this, daily weather changes should explain variation only in the uninformed order-flow.

The daily weather data from January 1, 1995 through the end of 2002 contain the following variables: hours of sunshine per day, precipitation in millimeters, cloud cover at noon (a scale from 1 to 8), and the noon temperature in Celsius degrees. These variables measured at the weather measurement station closest to the Helsinki Stock Exchange in Helsinki, Finland.[16] We deseasonalize these variables following the procedure that is used in Hirshleifer and Shumway (2003): we subtract from the daily value of the variable its average over the corresponding month. Out of these variables, we choose to employ only sunshine, because,

[15]Some papers argue that weather may affect other market variables than trading activity. Hirshleifer and Shumway (2003) find a positive relationship between weather and stock returns in a study of 26 stock markets around the world. Kamstra, Kramer, and Levi (2003) find evidence of a "winter blues," or a Seasonal Affective Disorder (SAD) effect on returns in various countries. By contrast, Goetzmann and Zhu (2005) argue that weather only affects stock returns because of its effect on spreads.

[16]Because approximately 25% of the Finnish population lives in the Metropolitan area, this weather directly applies to many investors. Moreover, because weather is spatially correlated, the weather in Helsinki is a good proxy for the weather even outside the capital.

consistent with Hirshleifer and Shumway (2003), we find that sunshine subsumes the other weather variables in explaining trading activity and volatility.

<< TABLE 2 ABOUT HERE >>

We run pooled panel regressions of log-arrival rates—e.g., the log-number of limit orders submitted by households—against the weather variables, controlling for the cross sectional variation with firm fixed effects. The results in Table 2, Panel A, indicate that weather influences household arrival rates considerably. For example, the standard deviation of the deseasonalized sunshine variable is 0.041. Thus, a one standard deviation shock to sunshine *decreases* household limit order arrival rates by 2.3% and market order arrival rates by 3.7%. Overall household trading activity decreases by 2.6%. These effects are both economically and statistically significant.

The result on institutional investors' arrival rates in Table 2, Panel B, offer further support for the hypothesized role of the weather variable: the behavior of investors, whether they are using market or limit orders, is almost completely unaffected by changes in weather. The results have neither statistical, nor economic significance. Panel C shows that when one aggregates the limit orders and market orders over both households and institutions, the results are statistically and economically significant, although not as strong as for households.

We use the Chicago Board Options Exchange Volatility Index (VIX) as the second instrumental variable. This index measures the market's current expectation of 30-day future stock market volatility. It is constructed using the implied volatilities of a wide range of S&P 500 index options, both puts and calls. The VIX is a widely used measure of market risk, and is sometimes called the "investor fear gauge." In our study we use the one-day lagged value of VIX as an instrument, as trading on the CBOE begins at the end of the trading day in Finland.

While not as strong an instrument as sunshine is for trading activity, we argue that VIX is a plausible instrument for volatility. For VIX to be a valid instrument, it should affect spreads only via its effect on volatility and trading activity, but not via any other channel. In particular, VIX should not affect the information ratio. This is plausible because in theoretical models, such as Roşu (2009b), the decision of becoming an informed trader depends on the costs of acquiring information, not on the fundamental volatility.

In Table 2, Panel D, we see that the effect of VIX on volatility is significant, both economically and statistically. A one standard deviation shock to VIX generates an increase in the stock midpoint volatility by 0.084%, which represents about one tenth of the standard deviation of the midpoint volatility ($\sigma = 0.87\%$).

From Panels A and B we note that a change in lagged VIX also affects the order submissions of households, but the order submissions of institutions do not seem to be affected. The only exception is institutional market order arrival rates: a one standard deviation increase in VIX leads to a 2.5% decrease in the number of institutional market orders. For households, a one standard deviation increase in VIX affects limit orders by more (5.4%) than it affects market orders (4%). This indicates that in times of high fundamental uncertainty (high VIX), all kinds of trading activity decrease, but households prefer to use market orders, while institutional traders prefer to place limit orders.

Panel D also implies that sunshine does not affect midpoint volatility. This provides an additional reason why weather is a good instrument: it strengthens the argument that weather only affects the traders' decision to participate in the market but it does not affect how they trade once they do come to the market. In particular, weather does not affect the fundamental volatility of the market. If it did, then it would also affect the midpoint volatility, as the two variables are directly related.[17]

Finally, to analyze the converse effect of trading activity on spreads, we introduce another instrumental variable: the number of modifications from limit orders to market orders when the limit order is at the bid or ask price (call this number MODIF_{ba}), normalized by the total number of modifications from limit to market orders (call this number MODIF). The identification assumption is that when trading activity is (exogenously) higher, both MODIF_{ba} and MODIF are higher, but they increase in the same proportion. This normalization ensures that the level of trading activity does not affect the limit-to-market order modification ratio. Moreover, if we count trading activity so that an order modification does not count as a

[17]Roşu (2009b) shows that the midpoint volatility is equal in steady state to the fundamental volatility.

new order, the limit-to-market order modification ratio is independent of trading activity. However, this modification ratio still affects spreads for a mechanical reason. If a modified order is the only order at the bid or ask, the disappearance of this order widens the bid-ask spread. Table 2, Panel E, shows that a one standard deviation increase in our instrument ($\sigma = 0.231$) leads to a 4.1% increase in the bid-ask spread. This suggests that our variable $\text{MODIF}_{ba}/\text{MODIF}$ is a good instrument: it is correlated with spreads, but it does not affect trading activity directly.

5 Empirical Results

5.1 Time Series Variation in Liquidity

This section examines whether exogenous variation in trading activity and volatility affect stock-specific liquidity. We study whether these variables *cause* liquidity by using the instrumental variables described in the previous section. To motivate the endogeneity issue, we first present an analysis that ignores the problem: we regress liquidity measures against contemporaneous trading activity, competition, and volatility. If the causality runs both ways—from activity to liquidity and from liquidity to activity—this is an endogenous regression. This is similar to the classical regression of quantity against price, where it is not clear whether we measure the demand curve or the supply curve. We then move to the main analysis, which uses instrumental variables.

<< TABLE 3 ABOUT HERE >>

Table 3 reports on a set of pooled panel OLS regressions of our liquidity measures against trading activity, competition, volatility, price, order autocorrelation, and order imbalance. Panels A, B, and C use total trading activity, while Panel D only uses household trading activity.

The endogenous-regression estimates for trading activity appear sensible: higher trading activity coincides with lower spreads and lower price impact measures. For example, these estimates suggest that a quarter standard deviation increase in log-trading activity decreases percentage bid-ask spreads by 3.64%. The median percentage bid-ask spread across the stocks in our sample is 2.27%. Then, a quarter standard deviation increase in trading activity narrows the spread to 2.19%.

However, in the case of trading competition, if we ignore endogeneity the results are counter-intuitive: when investors submit *more* limit orders, the bid-ask spread widens. This is almost certainly driven by reverse causality: when bid-ask spreads are wider, we expect traders to supply liquidity to the market by submitting more limit orders. If instead we exogenously added random limit orders to the market, the direct impact on the spreads would probably be the opposite. This is the endogeneity problem.

To make a comparison with the cross sectional study of Stoll (2000), we include more explanatory variables in the endogenous regressions from Table 3: midpoint price, order autocorrelation, and order imbalance. Stoll (2000) finds that spreads are higher in the cross section for stocks with lower trading volume, higher return volatility, lower price, and larger trading imbalances. The estimates in Table 3 confirm these results. We also find that wider spreads are associated with both higher competition and higher order-flow autocorrelation. Although we get similar results for the effective spreads, the coefficient estimates for trading competition, midpoint price, and order imbalance all change signs when the the price impact measure is the dependent variable. The results in Table 3 suggest that endogeneity issues are important, and instrumental variables are needed to unravel the endogeneity problem.

We also note that the t-statistics are much higher for trading activity, competition and volatility, than the t-statistics for midpoint price, order autocorrelation and order imbalance. Given the large number of observations, the lower statistical significance for the latter three variables suggests that it is reasonable to focus only on trading activity, competition, and volatility. Indeed, due to the difficulty in finding good instruments, we must restrict our attention to a few variables.

<< TABLE 4 ABOUT HERE >>

Table 4 focuses on trading activity and volatility, and uses weather and lagged CBOE Volatility Index (VIX) as instruments. The coefficient estimates are substantially higher than the estimates from the endogenous regression. Panel A shows that a quarter standard deviation *exogenous* increase in log-trading activity decreases percentage bid-ask spreads by 12.35%. (The decrease in bid-ask spreads is 13.5% if we also control for volatility and use the VIX as an instrument for it.) This means that the bid-ask spread of the median HEX stock would decrease from 2.27% to 1.99% if there was an exogenous shock to the arrival rates. A comparison between the endogenous estimates and the IV estimates suggests that the reverse causality is an issue. The endogenous-regression estimate is significantly lower, just 3.64%. By addressing the endogeneity problem, we identify a stronger effect from activity to spreads.

Similarly, a quarter standard deviation exogenous increase in volatility leads to a 14.1% increase in the bid-ask spread. This estimate is significantly larger than the endogenous-regression estimate of 5.48% in Table 3, Panel B. This result confirms the theoretical prediction that an exogenous increase in the fundamental volatility of an asset increases the bid-ask spread.

In addition to the bid-ask spread, we study the effect that trading activity and volatility have on the other two liquidity measures, the effective spread and price impact. The results on effective spreads closely parallel those for the bid-ask spread. Consistent with the previous results, we also find that higher trading activity lowers the price impact of a trade. The coefficient estimate from this IV regression (-0.406) is larger than the estimate from the OLS regression (-0.251). (Both regressions control for volatility.) Similar to the bid-ask spread estimates, this suggests that the "depth" in the limit order book, as measured by the price impact variable, invites more trading and leads to an endogeneity problem.

Numerically, a quarter standard deviation exogenous increase in trading activity decreases the price impact measure by 0.045, which represents about a quarter standard deviation change in the price impact. If we consider a stock with the median price impact in the sample, $\lambda = 2.063$, this corresponds to an implied market depth of $17.3 - e^{2.063} = 9.43$, and to a per-share price impact of $1/(e^{9.43} - 1) = 8.03 \times 10^{-5}$, i.e., to a price impact of 8.03% per 1000 shares. If price impact decreases by 0.045 to 2.018, this corresponds to a market depth of 9.78, and to a per-share price impact of 5.68×10^{-5}, i.e., to a price impact of 5.68% per 1000 shares.[18]

Another finding in Panel A of Table 4 is that the results on the price impact measure are stronger than the results on the bid-ask spread and the effective spread. This is intuitively appealing: we often use the bid-ask spread because it is easy to compute, but it is probably a noisier estimate of the market depth component of liquidity.

Panel B studies the effect of the bid-ask spread on trading activity. Biais, Hillion, and Spatt (1995) show that in the Paris Bourse a smaller bid-ask spread attracts more market orders but fewer limit orders, so it is not clear what the net effect is on overall trading activity. We find that indeed a lower bid-ask spread *causes* more trading activity. If the log-bid-ask spread decreased exogenously by a quarter standard deviation, then trading activity would increase by about 29%.

5.2 Robustness

We now verify the robustness of our results by performing the same regressions as above but splitting the sample in various ways. First, we split the sample in two time periods of equal length. The estimates for the early- and the late-sample are very similar to the full-sample estimates. Second, we examine below-median and above-median market capitalization firms separately and find the same results for both small and large firms. We also confirm that our results are not specific to the procedure we use to deseasonalize the weather variable. For example, we used an alternative procedure in which, instead of subtracting the monthly average of sunshine hours per day, we subtract the number of hours of daylight based on Finland's latitude.

If we introduce a dummy for the summer months (June, July, and August), the slope coefficient estimate for this variable is statistically insignificant and leaves the results in Table 2, Panels A, B, and C unchanged. The slope coefficient estimate for an analogous winter dummy

[18] These price impact estimates are high because we compute them for hypothetical trades. A hypothetical trade is sometimes "too large" given the number of shares available in the limit order book. For more details, see the definition of the price impact measure from Section 3.3.

is statistically significant and positive, but the inclusion of this variable also strengthens the link between deseasonalized sunshine and trading activity.

Despite these robustness tests, it remains a concern that there is some omitted seasonal effect that we do not control for. Under a view that our results are driven by such an omitted effect, trading activity is uncorrelated with weather. We rule out this possibility by studying the correlation between trading activity and weather within the same day. The benefit of our weather data is that we have data for 23 weather stations across Finland. Additionally, the complete trading and ownership records indicate the place of residence of each individual. If owners of two companies reside in different parts of the country, they are exposed to different weather conditions.

We construct a company-specific weather variable and find that trading activity in a stock is correlated with the same-company weather variable. First, we assign to each individual investor the weather station closest to the place of residence of that investor. Second, for each company, we set its corresponding weather variable equal to the value obtained from the station closest to most company shareholders.

Instead of running a separate cross sectional regression each day, we run a pooled panel regression with time fixed effects. The dependent variable is the log-trading activity in each stock and the independent variable is the company-specific weather variable. In this regression, we use the same deseasonalized sunshine variable as in the main text.[19] The estimates from these regressions indicate that weather is reliably correlated with trading activity in the company stock. The full-sample coefficient estimate for the company-specific weather variable is -0.019 with a t-statistic of -3.63. This estimate indicates that if the company-specific sunshine for company A is one standard deviation higher than for company B, the trading activity in company A is 1.5% lower than for company B.[20] This estimate is close to the time series estimate of 1.7% in the main body of the paper.

[19]The deseasonalization procedure is inconsequential because of the presence of the time fixed effects.

[20]We compute the standard deviation of the company-specific sunshine for each day and then average these estimates over time.

6 Conclusions

We examine the effect of trading activity and volatility on the time series variation in liquidity in a limit order market. As predicted by recent theoretical work, we find that these variables are important causes of liquidity. The causal statement is important because all these variables are determined endogenously. For example, although higher trading activity may reduce spreads, smaller spreads by themselves may also invite more trading activity. Thus, an OLS regression of one on the other fails to identify the causal link between them. The resulting problem is akin to the lack of identification in a regression of prices on quantity. By using instrumental variables for both trading activity and bid-ask spreads, we show that the causal relation works in both directions.

By using weather and the lagged CBOE Volatility Index as instruments, we find that higher trading activity and lower volatility generate lower bid-ask spreads and instantaneous price impact measures. These effects are both statistically and economically significant. For example, in the first stage of the instrumental variables regression, a one standard deviation shock to the weather variable decreases limit order arrival rates by 1.5%, market order arrival rates by 2.3%, and all trading activity by 1.7%. In the second stage of the IV regression, we find that even a quarter standard deviation *exogenous* shock to trading activity decreases percentage bid-ask spreads by 12.35%. This estimate is more than three times as large as the estimate obtained from an OLS regression (3.64%). This effect is economically significant. For example, the median percentage bid-ask spread across the stocks in our sample is 2.27%. Then, a one-quarter standard deviation exogenous shock to trading activity would cause this spread to narrow to 1.99%.

We restrict ourselves to analyzing the effect of noise trading activity and fundamental volatility on liquidity. Although we also would like to examine the role of asymmetric information on liquidity, we are constrained by the lack of a proper instrument. Moreover, we would like to be able to separate the common component of the fundamental volatility (instrumented by VIX) and the firm-specific, idiosyncratic component. In this paper we assume that both of these components affect liquidity in the same way, but we would like to examine whether this is true in the data.

References

Acharya, V. and L. Pedersen (2005). Asset pricing with liquidity risk. *Journal of Financial Economics 77*, 375–410.

Admati, A. R. and P. Pfleiderer (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies 1*, 3–40.

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

Black, F. (1971). Towards a fully automated stock exchange. *Financial Analysts Journal 27*, 28–33, 44.

Bollen, N., T. Smith, and R. Whaley (2004). Modeling the bid-ask spread: Measuring the inventory-holding premium. *Journal of Financial Economics 72*, 97–141.

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

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

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

Cohen, K. J., S. F. Maier, R. A. Schwartz, and D. K. Whitcomb (1981). Transaction costs, order placement strategy, and existence of the bid-ask spread. *Journal of Political Economy 89*, 287–305.

Demsetz, H. (1968). The cost of transacting. *Quarterly Journal of Economics 82*, 33–53.

Foucault, T. (1999). Order flow composition and trading costs in a dynamic limit order market. *Journal of Financial Markets 2*, 99–134.

Foucault, T., O. Kadan, and E. Kandel (2005). Limit order book as a market for liquidity. *Review of Financial Studies 18*, 1171–1217.

George, T., G. Kaul, and M. Nimalendran (1991). Estimation of the bid-ask spread and its components: A new approach. *Review of Financial Studies 4*, 623–656.

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

Glosten, L. and P. Milgrom (1985). Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 13*, 71–100.

Goettler, R. L., C. A. Parlour, and U. Rajan (2005). Equilibrium in a dynamic limit order market. *Journal of Finance 60*, 2149–2192.

Goetzmann, W. N. and N. Zhu (2005). Rain or shine: Where is the weather effect? *European Financial Management 11*, 559–578.

Grinblatt, M. and M. Keloharju (2000). The investment behavior and performance of various investor-types: A study of Finland's unique data set. *Journal of Financial Economics 55*, 43–67.

Handa, P. and R. A. Schwartz (1996). Limit order trading. *Journal of Finance 51*, 1835–1861.

Hirshleifer, D. and T. Shumway (2003). Good day sunshine: Stock returns and the weather. *Journal of Finance 58*, 1009–1032.

Huang, R. D. and H. R. Stoll (1997). The components of the bid-ask spread: A general approach. *Review of Financial Studies 10*, 995–1034.

Kamstra, M., L. A. Kramer, and M. D. Levi (2003). Winter blues: A sad stock market cycle. *American Economic Review 93*, 324–343.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica 53*, 1315–1336.

Linnainmaa, J. T. (2009). Do limit orders alter inferences about investor performance and behavior? *Journal of Finance*, forthcoming.

Parlour, C. A. (1998). Price dynamics in limit order markets. *Review of Financial Studies 11*, 789–816.

Pástor, Ľ. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111*, 642–685.

Ranaldo, A. (2004). Order aggressiveness in limit order book markets. *Journal of Financial Markets 7*, 53–74.

Roşu, I. (2009a). A dynamic model of the limit order book. *Review of Financial Studies*, forthcoming.

Roşu, I. (2009b). Liquidity and information in limit order markets. University of Chicago, working paper.

Sandås, P. (2001). Adverse selection and competitive market making: Empirical evidence from a limit order market. *Review of Financial Studies 14*, 703–734.

Seppi, D. J. (1997). Liquidity provision with limit orders and a strategic specialist. *Review of Financial Studies 10*, 103–150.

Stoll, H. (2003). Market microstructure. In G. Constantinides, M. Harris, and R. Stulz (Eds.), *Handbook of the Economics of Finance*. Elsevier.

Stoll, H. R. (2000). Friction. *Journal of Finance 55*, 1479–1514.

Table 1: Summary Statistics for Liquidity Measures, Determinants of Liquidity, and Instrumental Variables

We use intraday limit order book data from the Helsinki Stock Exchange from July 10, 2000 through October 23, 2001 (327 trading days) to construct three daily liquidity measures: bid-ask spread, effective spread, and (instantaneous) price impact. This table shows summary statistics for these liquidity measures, as well as for selected determinants of liquidity: ln(trading activity), competition, midpoint price, midpoint volatility (computed as the daily volatility of ln(midpoint price)), order autocorrelation and order imbalance. Panel A displays summary statistics for the 97 stocks selected for the sample. (We restrict the sample to stocks that average at least five marketable limit orders a day.) We first compute stock-specific averages for each measure and then report percentiles over these averages. We compute adjusted standard deviations for the stock-specific variables by first removing from each variable stock-specific means. The last column in Panel A reports the amount of cross sectional variation in each of the liquidity measures: the statistic is the adjusted R^2 from regressing the liquidity measure against firm fixed effects. Panel B uses the panel data to compute adjusted correlations between various variables. Each variable is first demeaned by subtracting the stock-specific averages.

Panel A: Summary Statistics

Market Quality Measures	Percentile 5%	25%	Median	75%	95%	Mean	Adj. Std. Dev.	ANOVA R^2
ln(Bid-Ask Spread)	−5.024	−4.438	−3.940	−3.452	−3.027	−3.967	0.513	63.94%
ln(Effective Spread)	−4.711	−4.110	−3.672	−3.273	−2.754	−3.714	0.489	64.13%
Bid-Ask Spread	0.72%	1.40%	2.27%	3.54%	5.87%	2.75%	1.59%	48.96%
Effective Spread	0.98%	1.83%	2.92%	4.20%	6.79%	3.38%	1.81%	49.90%
Price Impact	1.317	1.715	2.063	2.208	2.521	1.976	0.197	75.65%
ln(Activity)	3.125	3.722	4.157	4.871	6.263	4.398	0.500	81.61%
ln(Household Activity)	2.693	3.354	3.827	4.485	5.833	3.982	0.492	78.60%
ln(Institutional Activity)	1.094	2.126	2.703	3.787	5.588	2.994	0.747	77.96%
Competition	2.373	3.661	4.618	5.449	6.827	4.586	2.535	21.93%
Midpoint Volatility	0.58%	0.73%	0.98%	1.41%	1.78%	1.08%	0.87%	17.42%
Midpoint Price	1.072	3.898	7.880	18.511	35.656	12.192	5.102	84.63%
Order Autocorrelation	0.271	0.295	0.303	0.315	0.335	0.300	0.099	2.97%
Order Imbalance	−0.246	−0.118	−0.019	0.033	0.127	−0.045	0.479	5.81%
Deseasonalized Sunshine	−0.069	−0.024	−0.008	0.033	0.064	0.000	0.041	
CBOE Volatility Index	18.260	20.815	23.690	27.430	34.110	24.555	4.940	
ln(Norm. Limit-to-Market Order Modifications)	−0.471	−0.239	−0.116	−0.065	−0.016	−0.165	0.231	28.65%

Panel B: Adjusted Correlations

	Bid-Ask Spread	Effective Spread	Price Impact	ln(Act.)	ln(H-hold Act.)	ln(Inst. Act.)	Compet.	Midp. Volat.	Midp. Price	Order Autoc.
Bid-Ask Spread	1.000									
Effective Spread	0.882	1.000								
Price Impact	0.380	0.395	1.000							
ln(Activity)	−0.298	−0.323	−0.509	1.000						
ln(H-hold Activity)	−0.286	−0.287	−0.544	0.888	1.000					
ln(Inst. Activity)	−0.216	−0.264	−0.283	0.766	0.490	1.000				
Competition	0.166	0.152	−0.056	−0.237	−0.150	−0.320	1.000			
Midpoint Volatility	0.237	0.255	0.133	0.304	0.254	0.280	−0.172	1.000		
Midpoint Price	−0.104	−0.076	−0.012	0.060	0.026	0.079	0.002	−0.099	1.000	
Order Autocorrelation	0.015	0.004	−0.067	0.136	0.138	0.097	0.098	0.060	0.060	1.000
Order Imbalance	0.006	−0.006	−0.040	0.060	0.040	0.052	−0.111	0.009	−0.014	−0.040

Table 2: OLS Panel Regressions of Log-Arrival Rates and Volatility against Instrumental Variables: Weather and the CBOE Volatility Index (VIX)

This table uses daily data from the Helsinki Stock Exchange from July 10, 2000 through October 23, 2001 (327 trading days) to examine how the order submission strategies for households and institutions depend on the instrumental variables: sunshine, and the one day lagged Chicago Board Options Exchange Volatility Index (VIX). Panels A, B, and C report the results from a panel OLS regression of the log-number of orders against the instrumental variables. We use daily data for all stocks traded during the sample period. This regression is estimated separately for limit orders, market orders, and all orders. The daily weather variable represents the number of hours of sunshine per day, measured at the weather measurement station closest to the Helsinki Stock Exchange in Helsinki, Finland. We deseasonalize this variable following the procedure that is used in Hirshleifer and Shumway (2003). The reported sunshine coefficients are divided by 100. Panel D examines how the midpoint volatility (standard deviation of ln(midpoint price)) is affected by the weather and the lagged VIX. Panel E reports the results of an OLS regression of ln(bid-ask spread) on the logarithm of another instrumental variable: the number of modifications from limit orders to market orders when the limit order is at the bid or ask price, normalized by the total number of modifications from limit to market orders. The panel regressions in Panels A through E include firm fixed effects. T-statistics are reported in parentheses, and are obtained from standard errors clustered by firm.

Panel A: OLS Regressions of ln(Number of Household Orders) on Sunshine and Lagged CBOE Volatility Index (VIX), July 10, 2000–October 23, 2001

	Deseasonalized Sunshine	Lagged VIX	Adjusted R^2	Number of Obs.
ln(Household Limit Orders)	-0.569^{***}		0.26%	28,701
	(-9.46)			
	-0.526^{***}	-0.011^{***}	1.62%	27,734
	(-9.34)	(-5.39)		
ln(Household Market Orders)	-0.908^{***}		0.24%	28,701
	(-9.53)			
	-0.844^{***}	-0.008^{***}	0.49%	27,734
	(-9.22)	(-3.12)		
ln(Household Activity)	-0.629^{***}		0.28%	28,701
	(-10.17)			
	-0.583^{***}	-0.011^{***}	1.48%	27,734
	(-10.12)	(-5.22)		

Panel B: OLS Regressions of ln(Number of Institutional Orders) on Sunshine and Lagged CBOE Volatility Index (VIX), July 10, 2000–October 23, 2001

	Deseasonalized Sunshine	Lagged VIX	Adjusted R^2	Number of Obs.
ln(Institutional Limit Orders)	0.014		0.00%	28,701
	(0.13)			
	0.044	-0.001	0.01%	27,734
	(0.40)	(-0.50)		
ln(Institutional Market Orders)	-0.184		0.01%	28,701
	(-1.45)			
	-0.185	-0.005^{**}	0.12%	27,734
	(-1.51)	(-2.55)		
ln(Institutional Activity)	-0.058		0.00%	28,701
	(-0.50)			
	-0.040	-0.003	0.05%	27,734
	(-0.35)	(-1.30)		

Panel C: OLS Regressions of ln(Number of Orders) on Sunshine and Lagged CBOE Volatility Index (VIX), July 10, 2000–October 23, 2001

	Deseasonalized Sunshine	Lagged VIX	Adjusted R^2	Number of Obs.
ln(Limit Orders)	-0.373^{***}		0.11%	28,701
	(-5.66)			
	-0.345^{***}	-0.007^{***}	0.72%	27,734
	(-5.69)	(-3.80)		
ln(Market Orders)	-0.549^{***}		0.10%	28,701
	(-5.30)			
	-0.525^{***}	-0.007^{***}	0.36%	27,734
	(-5.46)	(-3.25)		
ln(Activity)	-0.413^{***}		0.12%	28,701
	(-5.81)			
	-0.386^{***}	-0.008^{***}	0.68%	27,734
	(-5.92)	(-3.83)		

Panel D: OLS Regressions of Volatility on Sunshine and the Lagged CBOE Volatility Index (VIX), July 10, 2000–October 23, 2001

	Deseasonalized Sunshine	Lagged VIX	Adjusted R^2	Number of Obs.
Midpoint Volatility	-0.019%	$0.017\%^{***}$	0.93%	27,491
	(-0.15)	(9.72)		

Panel E: OLS Regression of ln(Bid-Ask Spread) on ln(Normalized Limit-to-Market Order Changes)

	ln(Norm. Limit-to-Market Order Changes)	Adjusted R^2	Number of Obs.
ln(Bid-Ask Spread)	0.177^{***}	0.63%	28,669
	(8.40)		

Table 3: Endogenous Regressions of Liquidity Measures on Determinants of Liquidity

We use intraday limit order book data from the Helsinki Stock Exchange in Finland from July 10, 2000 through October 23, 2001 (327 trading days) to construct three liquidity measures: bid-ask spread, effective spread, and (instantaneous) price impact. This table runs pooled panel regressions of these liquidity measures against determinants of liquidity: ln(trading activity) or ln(household trading activity), as well as competition, midpoint price, midpoint volatility, order autocorrelation, and order imbalance. Trading activity is measured as the total number of market and limit order submissions. Trading competition is measured as the ratio of the numbers of limit orders and market orders. Midpoint price is the average between the bid and ask prices. Midpoint volatility is computed as the standard deviation of ln(midpoint price). Order autocorrelation is an average over the probability of a certain order type (buy market order, buy limit order, sell market order, sell limit order), conditional on the same type of order occurring immediately before. Order imbalance is the difference between the number of buy market orders and sell market orders, normalized by their sum. We demean all these variables by subtracting off stock-specific averages, to ensure that the R^2s from the regressions do not reflect variation attributable to differences in means. T-statistics are reported in parentheses, and are obtained from standard errors clustered by firm.

Panel A: OLS Regressions of ln(Bid-Ask Spread) on Determinants of Liquidity

ln(Activity)	Competition	Midpoint Volatility	Midpoint Price	Order Autocorr.	Order Imbalance	Adj. R^2	Number of Obs.
-0.291^{***}						7.88%	28,669
(-16.62)							
-0.264^{***}	0.024^{***}					9.24%	28,669
(-15.11)	(10.96)						
	0.046^{***}	17.957^{***}				12.31%	28,443
	(16.31)	(13.02)					
-0.462^{***}		25.196^{***}				24.33%	28,443
(-31.08)		(17.92)					
-0.430^{***}	0.031^{***}	26.025^{***}				26.60%	28,443
(-31.85)	(17.84)	(18.53)					
-0.431^{***}	0.032^{***}	30.914^{***}	-0.012^{***}	0.252^{***}	0.044^{***}	29.64%	26,943
(-31.55)	(18.25)	(20.09)	(-7.14)	(7.02)	(3.81)		

Panel B: OLS Regressions of ln(Effective Spread) on Determinants of Liquidity

ln(Activity)	Competition	Midpoint Volatility	Midpoint Price	Order Autocorr.	Order Imbalance	Adj. R^2	Number of Obs.
-0.315^{***}						9.78%	28,495
(-23.76)							
-0.292^{***}	0.019^{***}					10.73%	28,495
(-21.47)	(8.66)						
	0.042^{***}	16.860^{***}				11.58%	28,337
	(14.99)	(12.85)					
-0.480^{***}		24.666^{***}				26.67%	28,337
(-36.72)		(18.08)					
-0.452^{***}	0.026^{***}	25.351^{***}				28.41%	28,337
(-36.04)	(14.60)	(18.60)					
-0.457^{***}	0.027^{***}	30.871^{***}	-0.007^{***}	0.192^{***}	0.029^{***}	30.66%	28,337
(-34.24)	(15.43)	(20.41)	(-3.87)	(5.13)	(2.77)		

Panel C: OLS Regressions of Price Impact on Determinants of Liquidity

ln(Activity)	Competition	Midpoint Volatility	Midpoint Price	Order Autocorr.	Order Imbalance	Adj. R^2	Number of Obs.
-0.208^{***}						27.82%	28,701
(-17.63)							
-0.224^{***}	-0.015^{***}					31.14%	28,701
(-19.57)	(-10.20)						
	-0.003^{**}	2.301^{***}				1.35%	28,443
	(-2.31)	(7.29)					
-0.251^{***}		7.565^{***}				35.68%	28,443
(-26.81)		(16.73)					
-0.263^{***}	-0.012^{***}	7.245^{***}				38.01%	28,443
(-27.00)	(-8.83)	(16.19)					
-0.271^{***}	-0.012^{***}	8.376^{***}	0.002^{***}	0.028^{*}	-0.009^{***}	38.16%	26,943
(-26.41)	(-8.52)	(18.28)	(3.67)	(1.76)	(-2.91)		

Panel D: OLS Regressions of Liquidity Measures on ln(Household Activity) and Volatility

Dependent Variable	ln(Household Activity)	Midpoint Volatility	Adjusted R^2	Number of Obs.
ln(Bid-Ask Spread)	-0.293^{***}		7.72%	28,669
	(-15.60)			
	-0.439^{***}	23.588^{***}	22.57%	28,443
	(-27.17)	(18.15)		
ln(Effective Spread)	-0.290^{***}		7.97%	28,495
	(-17.82)			
	-0.425^{***}	22.431^{***}	22.34%	28,337
	(-25.55)	(18.09)		
Price Impact	-0.224^{***}		31.19%	28,701
	(-18.96)			
	-0.261^{***}	7.087^{***}	38.19%	28,443
	(-25.39)	(16.53)		

Panel E: OLS Regression of ln(Single-Counted Activity) on ln(Bid-Ask Spread)

	ln(Bid-Ask Spread)	Adjusted R^2	Number of Obs.
ln(Activity)	-0.268^{***}	8.35%	28,669
	(-11.06)		

Table 4: Instrumental Variables Regressions of Liquidity Measures on Trading Activity and Volatility

We use intraday limit order book data from the Helsinki Stock Exchange in Finland from July 10, 2000 through October 23, 2001 (327 trading days) to construct three liquidity measures: bid-ask spread, effective spread, and (instantaneous) price impact. This table runs pooled panel regressions of these liquidity measures against *instrumented* determinants of liquidity: ln(trading activity) or ln(household trading activity), as well as midpoint volatility. Trading activity is measured as the total number of market and limit order submissions. Midpoint price is the average between the bid and ask prices. Midpoint volatility is computed as the standard deviation of ln(midpoint price). We demean all these variables by subtracting off stock-specific averages, to ensure that the R^2s from the regressions do not reflect variation attributable to differences in means. T-statistics are reported in parentheses, and are obtained from standard errors clustered by firm. Panel A uses trading activity (both total activity and household activity), and volatility as the independent variables and uses weather (measured by the deseasonalized number of hours of sunshine per day) and the lagged CBOE Volatility Index (VIX) as instrumental variables. Panel B regresses trading activity on log-bid-ask spreads, where activity is obtained by counting modified limit orders as a single order. It uses as instrumental variable the number of modifications from limit orders to market orders when the limit order is at the bid or ask price, and normalizes it by the total number of modifications from limit to market orders.

Panel A: IV Regressions of Liquidity Measures on ln(Activity) and Volatility

Dependent Variable	ln(Activity)	ln(Household Activity)	Midpoint Volatility	Number of Obs.	Instruments
ln(Bid-Ask Spread)	−0.988*** (−4.06)			28,669	Sunshine
	−1.080*** (−5.05)		64.866*** (5.50)	27,491	Sunshine, Lagged VIX
		−0.650*** (−4.63)		28,669	Sunshine
		−0.724*** (−5.28)	65.092*** (5.84)	27,491	Sunshine, Lagged VIX
ln(Effective Spread)	−0.824*** (−3.77)			28,495	Sunshine
	−0.975*** (−5.21)		50.368*** (4.38)	27,388	Sunshine, Lagged VIX
		−0.542*** (−4.14)		28,495	Sunshine
		−0.639*** (−5.34)	51.003*** (4.63)	27,388	Sunshine, Lagged VIX
Price Impact	−0.363*** (−4.61)			28,701	Sunshine
	−0.406*** (−4.23)		34.818*** (6.07)	25,891	Sunshine, Lagged VIX
		−0.238*** (−5.88)		28,701	Sunshine
		−0.272*** (−4.15)	34.903*** (5.83)	27,491	Sunshine, Lagged VIX

Panel B: IV Regression of ln(Single-Counted Activity) on ln(Bid-Ask Spread)

	ln(Bid-Ask Spread)	Number of Obs.	Instruments
ln(Activity)	−2.265*** (−9.10)	28,669	Norm. Limit-to-Market Order Changes