Pre-trade Transparency, Market Quality, and Informed Trading

ARTO THURLIN*

Hanken School of Economics

This draft: January 15, 2009

Abstract

We study the effects of a change from a pre-trade transparent limit order book to an anonymous electronic limit order book. We use the Probability of Informed Trading (PIN) as a proxy for the participation rate of informed traders to test a theoretical prediction of Foucault, Moinas, and Theissen (2007). However, we do not find significant explanatory power for the PIN measure in explaining bid-ask spreads after a change to anonymous trading. More generally, we do not find unambiguous evidence of an improvement in market quality, measured by bid-ask spreads, intraday volatility, and trading volume. We also find that there is no substantial change in upstairs trading.

Keywords: Market transparency, informed trading, event study, upstairs market.

JEL Classification: G12, G14, G15.

*Corresponding author: Hanken School of Economics, PO Box 479, 00101 Helsinki, Finland.
E-mail: arto.thurlin[at]hanken.fi, Telephone: +358 40 352 1243, Fax: +358 9 431 33393.
We would like to thank Eva Liljeblom, Anders Ekholm, Mats Hansson, Thierry Foucault, Albert Menkveld, Peter Swan, Joakim Westerholm, Tommi Vuorenmaa, Aku Penttinen, seminar participants at Hanken School of Economics, the GSF Summer Research Workshop, and the Universität Mannheim Summer School for comments and suggestions on earlier versions of this paper. All errors are of course ours. In addition, we thank the OMX Group for providing data. We gratefully acknowledge financial support from the Hanken Foundation, the Finnish Foundation for Advancement of Securities Markets, Säästöpankkien tutkimussäätiö, and Osuuspankkien tutkimussäätiö. Some of this research was conducted at the Australian School of Business, UNSW, Sydney.

Pre-trade Transparency, Market Quality, and Informed Trading

Abstract

We study the effects of a change from a pre-trade transparent limit order book to an anonymous electronic limit order book. We use the Probability of Informed Trading (PIN) as a proxy for the participation rate of informed traders to test a theoretical prediction of Foucault, Moinas, and Theissen (2007). However, we do not find significant explanatory power for the PIN measure in explaining bid-ask spreads after a change to anonymous trading. More generally, we do not find unambiguous evidence of an improvement in market quality, measured by bid-ask spreads, intraday volatility, and trading volume. We also find that there is no substantial change in upstairs trading.

Keywords: Market transparency, informed trading, event study.

JEL Classification: G12, G14, G15.

I. Introduction

Should an exchange be as transparent as possible, revealing prices, market depth, and the identities of traders posting bid and ask prices to all market participants? And if so, who gains and who loses if a transparent market is made less transparent? The question is timely, and it affects the ongoing competition and consolidation in global equities markets. In order to accommodate the needs of different kinds of investors, several exchanges, e.g. in Sydney, Seoul, Paris, Tokyo, and Helsinki have made changes in market transparency during the last few years[1].

We study the change from a transparent market to an anonymous market at the OMX Helsinki Stock Exchange in March 2006. Using the Probability of Informed Trading (PIN) as a proxy for the participation rate of informed traders, we test a theoretical prediction by Foucault, Moinas, and Theissen (2007). According to their theory, a higher participation rate of informed traders is correlated with a higher bid-ask spread. However, in a cross-sectional regression we do not find evidence of a significant explanatory power for the PIN variable.

We also study whether market quality has improved, measured in a smaller bid-ask spread, higher trading volume, or diminished intraday volatility. The evidence is mixed, with an unchanged average bid-ask spread, higher trading volume, and higher volatility. As a comparison to the electronic limit order book, we study the upstairs market, which is unaffected by the changes in the limit order book. We find no change in the numbers of upstairs trades, or internalization rates.

[1] Euronext Paris, Tokyo Stock Exchange, and the Australian Stock Exchange removed the display of broker identifiers from the limit order book on April 23, 2001, June 30, 2003, and November 28, 2005, respectively. The Korea Stock Exchange made the opposite move, introducing broker IDs on October 25, 1999. OMX Helsinki, the object of the present study, removed the display of broker IDs on March 13, 2006.

Theoretical research reaches mixed conclusions about the effects of transparency. Papers supporting the view that increased transparency enhances liquidity include Admati and Pfleiderer (1991), Pagano and Roell (1996), and Baruch (2005). The most common explanation is the mitigated effect of information asymmetry (see, e.g. Chowdhry and Nanda (1991), Madhavan (1995), and Bloomfield and O'Hara (1999)). Pagano and Roell (1996) differentiate between uninformed and informed traders, and find that the former benefit from greater transparency.

Empirical research on the effects of market transparency was initially conducted as experiments, due to a lack of data. Experimental studies with human subjects in laboratory conditions are a promising avenue of research. Early papers, such as Bloomfield and O'Hara (1999, 2000), and Flood, Huisman, Koedijk, and Mahieu (1999), find fairly complex results regarding the effects of transparency. Bloomfield and O'Hara (1999) find that price discovery is more efficient in an opaque market, while other measures of the quality of the market, such as bid-ask spreads and trading volume deteriorate.

The results of the empirical study of the NASDAQ market by Harris and Schultz (1997) support market transparency. They find that the anonymous Small Order Execution System of the NASDAQ market makers show wider bid-ask spreads compared with the regular, non-anonymous dealer markets. Theissen (2003) finds a similar result for the German market. On the non-anonymous floor-based trading system of the Frankfurt Stock Exchange, specialists are able to credibly identify uninformed traders. When trading with an uninformed trader, the specialists are willing to offer price improvement. Transparency seems to enhance liquidity in this case.

Boehmer, Saar and Yu (2005) study the introduction of the electronic limit order market, OpenBook, at the NYSE. They find that this great increase in transparency to market participants outside the trading floor resulted in increased market depth and a reduced effective spread. Hendershott and Jones (2005) study a change to a less transparent trading system at Island, an Electronic Communications Network (ECN). Island stopped displaying the limit order book for the most liquid Exchange Traded Funds (ETFs), which decreased liquidity significantly. Madhavan, Porter, and Weaver (2005) examine an increase in pre-trade transparency on the Toronto Stock Exchange in 1990. In contrast to the above studies, they find that execution costs and volatility increased. The authors attribute the finding to the increased efficiency in order placement by market makers, and the decreased willingness of limit-order submitters to offer them free liquidity options. Other papers supporting less pre-trade transparency include Madhavan (1996), Garfinkel and Nimalendran (2003), and Foucault, Moinas, and Theissen (2007).

Most previous empirical research compares two exchanges with different characteristics. It is generally consistent with the notion that anonymity is associated with higher adverse selection costs. Huang and Stoll (1996) compare two markets, NYSE and NASDAQ, where a NYSE specialist can see the limit order book, and NASDAQ participants cannot. They find that spreads are generally higher on the more opaque NASDAQ. Chan and Lakonishok (1995) study institutional trading on NYSE and on NASDAQ. They find that smaller stocks have better execution on NASDAQ, and large stocks on NYSE.

Madhavan and Cheng (1997) find that, consistent with the model of Seppi (1990), the upstairs market is used by traders who can credibly signal that they trade for liquidity reasons. de Jong, Nijman, and Roell (1996) show that trades that are negotiated bilaterally (and thus non-anonymously) and are then executed through the Paris Bourse's CAC system have a lower price impact than regular CAC trades. Garfinkel and Nimalendran (2003) document that NYSE stocks exhibit larger increases in the bid–ask spread on insider trading days than NASDAQ stocks and conclude that the trading system of the NYSE is less anonymous. Grammig, Schiereck, and Theissen (2001) compare floor trading at the Frankfurt Stock Exchange with the electronic Xetra market. They find that informed traders tend to prefer the anonymous electronic market, whereas uninformed traders are drawn to the smaller adverse selection costs of the floor market. Hendershott and Jones (2005) study the decision by Island, an electronic communications network (ECN), to stop displaying its limit order book in its most liquid products to all market participants. They find that the changes resulted in higher trading costs, and lost market share for Island.

Another related paper is Simaan, Weaver, and Whitcomb (2003). They compare the quotation behavior of Nasdaq market makers on different trading platforms. They find that when dealers are able to quote anonymously, on ECNs (Electronic Communications Networks), the bid-ask spread narrows. This implies collusion among the dealers on the non-anonymous Nasdaq platform.

As originally noted by Copeland and Galai (1983), limit orders have option-like features. Sell (buy) limit orders are similar to a free call (put) option with a strike price equal to the price of the limit order. As option prices are dependent on volatility, limit order traders should also be able to use volatility information when placing their orders.

This treatment of limit orders as options, whose pricing is of course dependent on the volatility of the underlying security, is the starting point of the analysis by Foucault, Moinas, and Theissen (2007). They construct a three period model with several types of participants. In period 0 two kinds of limit order traders, value traders and pre-committed traders, post limit orders. Value traders post limit orders only if it is profitable to do so, whereas pre-committed traders are committed to buying or selling a given number of shares. Value traders are either informed or uninformed about future volatility. Being informed about future volatility in this case simply means knowing whether there is an information event in the next period. In period 1, speculators and liquidity traders trade using market orders. If there is an information event, a speculator arrives with a given probability, and buys or sells according to the observed price innovation. He is informed about expected volatility, and picks off buy or sell limit orders within the bounds given by the expected volatility. If there is no speculator, a liquidity trader is equally likely to buy or to sell, using a market order.

Foucault, Moinas, and Theissen (2007) model dealers whose limit orders may become stale and are subsequently picked off by faster speculators when volatility is high. The amount of informed trading is a key variable in this setting. In a transparent market with a low number of informed traders, uninformed traders learn from informed traders with volatility information. In this case bid-ask spreads widen in anticipation of adverse information events as dealers protect themselves. On the other hand, in an anonymous market spreads will narrow and the limit order book will become less informative. The bid-ask spread and its informativeness will both decline with the change to an anonymous market.

The opposite result obtains when there are lots of informed traders. In this case the uninformed traders are wary of trading with an informed trader. Thus the bid-ask spread will widen with anonymity. Intuitively this result can be understood as follows. In the case of few informed traders in the market, quotes are most likely posted by uninformed traders, and traders are not afraid to post aggressive limit orders, thus decreasing the bid-ask spread. When the number of informed traders increases, quotes are more likely to be informed. Anybody seeing a large bid-ask spread will be more cautious, and posts smaller orders, and at less aggressive prices. The bid-ask spread will either increase, or remain unchanged.

Another related paper is Rindi (2008), who studies the effect of pre-trade transparency on the informed traders' demand. She studies different types of markets, and concludes that factors such as the existence of potential insider information in a market can determine whether transparency is beneficial to a market. In a market where insider information is unlikely or non-existent, such as government bond and foreign exchange markets, and equity markets with strong insider regulations, less transparency would enhance market quality. The reason is that greater transparency reduces the incentive to acquire information. This reduces the number of informed traders, who are willing to accommodate liquidity demands from uninformed traders. The main difference between this model and the model of Foucault et al. (2007) is volatility, which is not analyzed by Rindi (2008).

Our main contribution is a novel way of using the estimated Probability of Informed Trading (PIN) as a proxy for the participation rate of informed traders. In this way we are able to test the theoretical model of Foucault, Moinas, and Theissen (2007). We first estimate the PIN measure for all stocks, for both the pre- and the post-change period. We then use the estimated values for PIN as an explanatory variable for the bid-ask spread, with a dummy variable to account for the change to an anonymous market. We do not find evidence PIN is a significant explanatory variable for the post-change spread. This result does not lend empirical support to the model proposed by Foucault et al. (2007).

We also contribute to the existing literature on the effects of a change in market transparency by studying a recent change in pre-trade transparency. Also, the object of our study, the OMX Helsinki market in 2006, has full post-trade transparency, i.e. the immediate disclosure to all market participants of the counterparties of each trade. Our empirical analysis indicates that the change to a pre-trade anonymous market in not as clearly beneficial to market quality as many previous studies suggest. We find a significant increase in intraday volatility, and no significant change in average bid-ask spreads. Results from a pooled regression show that we can attribute part of the changes in quoted spreads to the change to an anonymous trading system.

The changes made to the transparency of the limit order book do not affect the upstairs market. This market consists of brokers personally talking to other brokers, either within or outside the brokerage. The information about counterparties is in this way automatically present. We do not find any change in upstairs trading, which is our expected result.

The rest of this study is organized as follows. Chapter 2 discusses the data and the exchange studied. Chapter 3 presents the methodology for studying informed trading, including details of estimation of PIN. Chapter 4 discusses our empirical results, and Chapter 5 concludes.

II. Institutional details and the dataset

We study transparency of the order book at the OMX Helsinki Stock Exchange[2]. OMX Helsinki is an electronic limit order market. After an opening auction, continuous trading starts at 10 a.m., and ends in a closing auction starting at 6.20 p.m. Before the changes implemented on 13th March 2006, the trading system allowed all participants to see the identities of all brokers posting limit orders in the limit order book. After the change all identities are hidden. The only visible information is price and size for each limit order in the book. It is notable that these changes do not affect post-trade transparency. This means that the counterparties of each trade are known to all market participants immediately after a trade takes place.

We use a dataset provided by the Helsinki Stock Exchange. Our sample consists of the 35 most liquid stocks, measured by daily trading volume[3]. Our selection criteria are that the stocks be continuously listed during the entire sample period, and that the average daily trading volume be greater than one million euros.

Table 1 presents descriptive statistics of our sample companies. As is apparent both from the market capitalizations, and the trading volume statistics, our sample is very

---

[2] HEX, the Helsinki Stock Exchange, merged with the Swedish OMX Group in September 2003 to form OM HEX. OM HEX later changed their name to OMX Group. The company has since then become a major operator in the Nordic area, after merging with or acquiring the following marketplaces: Copenhagen, Stockholm, Helsinki, and Iceland in the Nordic market, and Tallinn, Riga, and Vilnius in the Baltic states. There is also an alternative exchange for small companies, First North. In 2007, NASDAQ acquired OMX Group, to form the NASDAQ OMX Group, Inc.

[3] There are a number of changes in the exchange listings during the sample period. Ahlstrom Corporation Oyj (AHL1V) was listed on March 14, 2006. Also, SanomaWSOY Oyj combined their two share series, SWSAV and SWSBV into one series, SWS1V, on April 10, 2006. We therefore exclude both these stocks are from our sample.

heterogeneous. Nokia is by far the largest and most liquid stock, both by market capitalization and trading volume. There is a second tier of companies, which could be said to include Fortum, Metso, Neste Oil, Outokumpu, Sampo, Stora Enso, and UPM-Kymmene. The rest of the sample companies are smaller and less liquid. All companies are liquid enough for the analysis of informed trading. The average daily number of trades is in the hundreds or thousands for almost all companies in our sample.

60 days is generally regarded as the minimum sample period for the reliable estimation of the Probability of Informed Trading, see e.g. Easley et al. (1997), Easley et al. (2005), and Aktas et al. (2007). On the other hand, most related event studies on exchange transparency use shorter sample periods, presumably to minimize the presence of other factors affecting the trading environment. With due consideration for these two factors, we use a total of six months of data in two sub-samples of equal size. There are 62 trading days before the changes of March 13, 2006, and 62 trading days after the changes. We exclude the day of the actual change, to give the market time to adjust to the new trading system. Our pre-change sample, which covers the period of non-anonymous trading, runs from December 13, 2005 to March 10, 2006. Our post-change (anonymous trading) sample period runs from March 14, 2006 to June 13, 2006.[4].

[4] Our sample periods of 62 + 62 days are longer than those used in many related studies, such as the sample periods of 20 + 20 trading days in Comerton-Forde et al. (2005), 25 + 29 days in Hendershott and Jones (2005), and 14 + 14 days in Foucault et al. (2007). Foucault et al. (2007) eliminate two weeks of trading around the event date. Comerton-Forde (2005) and Hendershott and Jones (2005) do not eliminate any days around the event dates.

III. Informed trading

A. The PIN (Probability of Informed Trading) model

The Probability of Informed Trading (PIN) measures information asymmetry between investors and traders in a stock. The model was introduced in a series of papers by Easley, Kiefer, O'Hara, and Paperman (1996), and Easley, Kiefer, O'Hara (1996, 1997). The model is based on the Glosten and Milgrom (1985) sequential trading model.

We follow the original PIN model very closely. Figure 1 presents an intuitive overview of the trading process. There are three types of traders in our model: informed traders, uninformed liquidity traders, and risk neutral market makers. At time zero an information event occurs. This event has either a positive ('high'), or negative ('low') effect on the value of the security. An informed trader is risk neutral, takes prices as given, and does not engage in any strategic behavior.

If an information event has occurred in period zero, in period one the informed trader acts upon this information, which only he or she possesses. If an informed trader has seen a high signal, he or she buys the stock if the current price is below the value given the signal; if the signal is low, he or she will sell if the quote is above the value given the signal. If there is no information event, an informed trader does not trade.

The uninformed trader's behavior is more complex. In most microstructure models, the presence of traders with better information dictates that an uninformed trader trading for speculative reasons would always be better off not trading at all. To avoid this no-trade equilibrium, at least some uninformed traders must transact for nonspeculative reasons such as liquidity needs or portfolio considerations. Since the uninformed traders do not have a particular reason to buy or sell, we make the customary assumption that half of them are buyers and the other half are sellers. We assume that when an uninformed trader checks the quote, the probability that he or she will trade is strictly positive.

Before trading starts each day, an information event occurs with probability α. There is thus a probability of (1- α) that there is no information, in which case only the uninformed traders are active in the market. If there is information, it is bad news ("low signal') with probability δ. The complementary event is that there is good news ('high signal") with probability (1- δ). In the former case, informed traders sell, in the latter, they buy.

During the actual trading process, traders arrive according to a Poisson process. The market maker is always ready to trade during the day, and posts buy and sell quotes accordingly. Orders from informed traders arrive according to a Poisson process, with the intensity parameter μ. Orders from uninformed traders arrive with parameters ε_s and ε_b for sell and buy orders.

Once we have estimated the above parameters, as discussed in Section III.B.2 below, we are in a position to calculate the Probability of Informed Trading, PIN, as

$$PIN = \frac{\alpha\mu}{\alpha\mu + \varepsilon_s + \varepsilon_b}, \tag{1}$$

where $\alpha\mu + \varepsilon_s + \varepsilon_b$ is the arrival rate of all orders, and $\alpha\mu$ is the arrival rate of informed orders. The probability of informed trading is thus the ratio of orders from informed traders to the total number of orders.

This model has been adapted to numerous uses, including a study of international analyst coverage in Easley, O'Hara, and Paperman (1998), stock splits in Easley, O'Hara, and Saar (2001), information risk and expected returns in Easley, Hvidkjaer, and O'Hara (2002), and the anonymity of the trading process in Grammig, Schiereck, and Theissen (2001).

B. Maximum likelihood estimation of the PIN model

In this section, we present the procedure for estimating the PIN model. We first need to calculate the input data for the estimation. These data are the numbers of buyer and seller initiated trades per trading day. We then use a numerical estimation procedure for the maximum likelihood estimation of the Probability of Informed Trading.

1. The Lee & Ready (1991) algorithm

Our starting point is ultra-high frequency transactions data. The first step is to determine the number of buyer initiated trades ('buys') and seller initiated trades ('sells') during each trading day. Since this information is not available in the data, it is necessary to employ one of several available methods of inferring the initiator of a trade. The three most common methods are the tick test, which uses changes in trade prices, the quote method, in which trade prices are compared to the prevailing bid and ask prices, and the Lee and Ready (1991) algorithm, which combines the tick and quote methods.[5]

The Lee and Ready (1991) algorithm has two steps in classifying trades. First, if the trading price is closer to the prevailing bid than the offer at the time of the trade, the trade is classified as a sell, and vice versa. Second, for trades that occur at the midpoint of the quote, a tick rule is used. According to this rule, if the last price change was positive (negative), the trade is a buy (sell). Following the findings of Bessembinder (2003), we do not use a time lag when matching quotes with trades. This means that for each trade we use the best bid and ask prices available at the time of the trade.[6]

[5] See Finucane (2000) and Bessembinder (2003) for a general discussion of the merits of the different methods, and a comprehensive discussion of their use in the literature; Ellis, Michaely, and O'Hara (2000) for tests on the three methods in the NASDAQ market; Chakrabarty et al. (2007) for a discussion on the merits of the methods in modern ECN trading systems; Savickas and Wilson (2003) discuss applications in options markets.

[6] See Bessembinder (2003) for a discussion of the appropriate lag length and alternative algorithms for signing trades. In their original paper, Lee and Ready (1991) recommend using quotes lagged by five seconds when assessing trades. This is done to allow for delays in trade reporting. Bessembinder (2003) also points out, that there are two separate issues at play: classifying trades into buys and sells, and assessing effective trade costs. In the former case, it is best to use a zero lag. In the latter case, it ca be argued that a lag should be used, since traders are concerned about the possibility of adverse price movements between the time of the trade decision and trade execution.

Boehmer, Grammig, and Theissen (2007) analyze the effects of inaccurately classifying trades as buys and sells. In general, the results of the Lee and Ready algorithm are known to be somewhat inaccurate. Based on both simulation results and empirical evidence, the authors argue that PIN estimates are downward biased, when trade classification is inaccurate. We have no reason to believe that our estimation results differ in accuracy from other existing literature in the field; however, the evidence presented by Boehmer et al. (2007) calls for some caution when interpreting our results.

2. The maximum likelihood model

The model is based on the assumption that order arrivals follow independent Poisson processes, as discussed in Section III.A above. These arrival intensities induce the following model for the total number of buy and sell trades during a single trading day:

$$L\left(\left(B,S\right)\middle|\theta\right)=\alpha\left(1-\delta\right)e^{-\left(\mu+\varepsilon_b+\varepsilon_b\right)T}\,\frac{\left(\mu+\varepsilon_b\right)^{B_i}\left(\varepsilon_s\right)^{S_i}}{B_i!S_i!}+ \\ \alpha\delta e^{-\left(\mu+\varepsilon_b+\varepsilon_b\right)}\,\frac{\left(\mu+\varepsilon_s\right)^{S_i}\left(\varepsilon_b\right)^{B_i}}{B_i!S_i!}+(1-\alpha)e^{-\left(\varepsilon_b+\varepsilon_b\right)}\,\frac{\left(\varepsilon_b\right)^{B_i}\left(\varepsilon_s\right)^{S_i}}{B_i!S_i!}, \tag{2}$$

where B_i and S_i are the total number of buy and sell trades on the day in question, and $\theta=\left(\mu,\varepsilon_b,\varepsilon_s,\alpha,\delta\right)$ is the parameter vector. This likelihood function emanates from the distributions of the trade outcomes, weighted by the probabilities of the three different types of trading days. The day can be entail positive news, with a probability of α(1 - δ), negative news, with a probability of αδ, or no news at all, with a probability of (1 - α).

The likelihood function over multiple trading days is the a product of the likelihood functions of a single day (Equation 2 above):

$$L\left(\theta\,|\,M\right)=\prod_{i=1}^{I}L\left(\theta\,|\,B_i,S_i\right), \tag{3}$$

Where B_i, and S_i are the numbers of buy and sell trades for day $i=1,\ldots,I$, and $M=((B_1,S_1),\ldots,(B_I,S_I))$ is the data set. Easley, Kiefer, and O’Hara (1997) test the assumption of independence among the trading days. They are not able to reject the independence assumption.

The direct computation of the maximum likelihood function may result in numerical overflow, since the values of *B!* and *S!* often become very large. We therefore perform the actual calculations using the following approximation, following Easley, Hvidkjaer, and O'Hara (2005). This approximation follows from the above models, after dropping a constant and rearranging terms.

$$
\begin{aligned}
L\left(\left(B_t,S_t\right)_{t=1}^{T}|\theta\right) &= \sum_{t=1}^{T}\left[-\varepsilon_b-\varepsilon_s+M_t\left(\ln x_b+\ln x_s\right)+B_t\ln\left(\mu+\varepsilon_b\right)+S_t\ln\left(\mu+\varepsilon_s\right)\right] \\
&+\sum_{t=1}^{T}\ln\left[\alpha\left(1-\delta\right)e^{-\mu}x_x^{S_t-M_t}x_b^{-M_t}+\alpha\delta e^{-\mu}x_b^{B_t-M_t}x_s^{-M_t}+\left(1-\alpha\right)x_s^{S_t-M_t}x_b^{B_t-M_t}\right]
\end{aligned} \tag{4}
$$

where $M_t=\min\left(B_t,S_t\right)+\max\left(B_t,S_t\right)/2$, $x_s=\dfrac{\varepsilon_s}{\mu+\varepsilon_s}$, and $x_b=\dfrac{\varepsilon_b}{\mu+\varepsilon_b}$. The advantages of this factorization are an increase in computing efficiency, and an avoidance of potential overflow problems, when the numbers of buys and sells are large[7].

IV. Empirical analysis

In this section we analyze the effects of the change to an anonymous market. First we perform a univariate analysis of the changes in market quality, measured in bid-ask spreads, trading volume and volatility. We then seek to explain these findings in a multivariate framework. Subsequently, we examine informed trading in the form of the PIN measure. We study changes in informed trading and use our estimates of PIN as an explanatory variable for

---

[7] In the estimation of this likelihood model, we employ the *fminsearch* function in Matlab, which is an implementation of the simplex algorithm. Also, the choice of the initial values is important. We therefore randomize our initial values. Out of the estimated maximum likelihood values we then pick the ones that give the greatest value for the likelihood function.

changes in market quality. We also study the upstairs market, in comparison to the continuous market, and conclude that no change has occurred there, as expected.

A. Univariate analysis

We start this paper by asking the normative question of whether an exchange should be transparent. Even if a categorical answer may not exist, we can still examine the effects of a change in transparency and anonymity, and see whether the market improves in quality. By a high market quality we mean a liquid market with a small cost of trading. In other words, we wish to observe a high trading volume, a low bid-ask spread, and low volatility. We measure the bid-ask spread in three ways: in euros, as a percentage spread, and as an effective spread. The bid-ask spread in euros is the difference between the ask price and the bid price. The percentage spread is defined as $\frac{ask - bid}{(ask + bid)/2} \times 100\%$, i.e. as the ratio of the bid-ask spread and the midquote.

We calculate the effective spread as follows:

$$Spread_{effective} = 2 * \left| P - m_{lag} \right|, \tag{5}$$

where P is the trade price, m_{lag} is the lagged midpoint of the best bid and ask prices. We use a lag of 0, 1, and 5 seconds. The motivation for estimating an effective spread is that it gives a better estimate of the real cost of trading than the quoted bid-ask spread. Originally proposed by Lee and Ready (1991), there is an extensive literature concerning the estimation and use of effective spreads. However, since a modern day electronic limit order market is very different from the NYSE in the early 1990s, the effective spread in our data differs from the quoted spread typically only when a large order executes at multiple prices, i.e. when the order depth at the best bid or ask price is insufficient.

Table 2 presents our univariate results of the measures of market quality, before and after the change to less transparent trading. The results show that there is no significant change in mean bid-ask spreads, measured both as percentage spreads, or in euros. A paired t-test fails to reject the null hypothesis of equal means. However, the median spreads are significantly smaller, according to the Wilcoxon measure. The economic significance of the change, 2.8% in the case of the spread measured in euros, and 3.3% in the percentage spread, is arguably quite small.

There are other significant changes in the market, however. There is no significant change in the average daily trading volume, measured in euros, or the number of shares traded. Traders seem to have shifted their behavior in that they have moved to trading in smaller lots. This can be seen in that the average daily number of trades increases, and the average trade size decreases. Both effects are significant, measured both in mean and median values. Last but not least, there is a significant rise in intraday volatility, measured as 30-minute price returns.

For market quality to have improved, bid-ask spreads would have to be lower, trading volume should be higher, and the volatility lower. The evidence is therefore ambiguous as to whether market quality has improved with this change to anonymous trading.

B. Multivariate results

Section IV.A above presents the changes in several variables that describe market quality. The effects described could of course be the result of other factors than the change in the trading system itself. In order to be able to attribute changes in these variables to this transition to an anonymous market in March 2006, we estimate the following regression, using a slightly modified version of the regression used by Foucault, Moinas, and Theissen (2007):

$$s_{i,t} = \gamma_0 + \gamma_1 \log\left(V_{i,t}\right) + \gamma_2 P_{i,t} + \gamma_3 \sigma_{i,t} + \gamma_4 D_t^{post} + \varepsilon_{i,t}, \tag{6}$$

where s is the spread of stock i at time t, V is the trading volume of the stock, P the stock price, σ is midquote return volatility. D^{post} is a dummy variable that takes on a value of

0 for the period before the changes (the transparent market), and a value of 1 after the changes (the opaque market)[8].

We first calculate all variables for each stock and each day. We then aggregate the variables over the duration of the two sample periods, producing two observations per stock, one for each sample period. We correct for potential autocorrelation by using Newey-West standard errors in calculating t-statistics. We perform the regression separately for our two sample periods, the pre and post change periods.

Our dependent variable is the quoted bid-ask spread, measured in two ways, in euros and as a percentage. The percentage spread is the difference of the bid and ask prices as a percentage of the stock price; the spread in euros is simply the difference between the two prices.

Table 3 presents our results. The main result is that when estimating regression model (5), the trading period dummy is highly significant. This means that some of the changes in market quality may be explained by the shift to anonymous trading. The period dummy has a positive sign (except for a non-significant negative sign in the case of the spread in euros with fixed effects), which is the expected result, since the dummy is zero for the pre-change period, and one for the post-change period.

All explanatory variables have the expected signs. Volatility has a positive estimated coefficient, as expected according to the Foucault, Moinas, and Theissen (2007) model. The reasoning is that uninformed limit order traders are afraid of being picked off by informed traders, and protect themselves by a greater bid-ask spread in the case of greater volatility. Trading volume has a negative sign, meaning that the greater the trading volume, the tighter is the bid-ask spread. This result is expected and in line with previous research. Price has a

[8] The only difference to the Foucault, Moinas, and Theissen (2007) regression is that they also include tick size as a variable. However, since the Helsinki market has a uniform tick size of EUR 0.01 for all stocks and all prices, we exclude this variable.

negative sign when using the percentage spread as the dependent variable, and a positive sign for the bid-ask spread in euros. This is also to be expected, since the greater the stock price, the greater the bid-ask spread, in euros, for the same relative (percentage) spread. Also, percentage spread can be expected to be a negative function of stock price, since the minimum tick size of 1 cent poses a natural lower limit for the spread.

As a robustness check, we perform all the above analyses using effective spreads as well as quoted spreads. As a further robustness check, we perform a company fixed-effects analysis. A fixed stock effect may be a source of correlation in the analysis. The results are similar to the results presented in Table 3, although the significance of the dummy variable is reduced. Both these robustness checks provide largely similar results to the base case presented in the paper. We therefore omit these results; they are available on request.

C. The upstairs market

The upstairs market[9] consists of prearranged trades, which are negotiated in person between brokers. They can occur "in-house", where the buyer and seller are both represented by the same brokerage, or between brokerages. The two main types of upstairs trades on the OMX Helsinki exchange are block trades and contract trades. A block trade is a large trade, defined as a certain minimum percentage of the free float of a stock, and a minimum size in value. These trades can be executed even outside the prevailing bid-ask spread, as long as they are within the limits of the lowest and highest prices for the day. Contract trades are trades that are not automatically matched in the limit order book, but previously agreed upon, either within the same brokerage, or between brokerages.

---

[9] There is a widely used distinction between the upstairs market, described above, and the downstairs market, which is the automatically matched continuous trading in the electronic limit order book.

Smith et al. (2001) find that the upstairs market of the Toronto Stock Exchange enables large non-information based trades to execute at a lower cost. Bessembinder and Venkataraman (2003) come to a similar conclusion for the Paris Bourse.

The changes in pre-trade transparency do not affect the upstairs market. The nature of the trading process is in this way the almost complete opposite to the recent proliferation of the so-called dark pools, where buyers and sellers are matched with no knowledge of the counterparties, and with no price discovery. For this reason it is interesting to find that there are no changes in the upstairs market. The internalization rates, i.e. the percentage of trades where a brokerage is able to find a counterparty for a trade among its own clients, do not change. The scale of internalization for upstairs trades is similar to the findings of Booth et al. (2000), at around 97.5%. Table 4 presents our results. We find that the internalization rate is consistently very high, both for block trades and for contract trades. A t-test of the internalization rates does not reject the null hypothesis of no change in the internalization rates.

D. The Probability of Informed Trading

Madhavan (1996) shows that increased transparency does not necessarily result in better liquidity. In a fully transparent market, which reveals the depth of the limit order book, the value of the stock is known with great precision. Such a market discourages uninformed (noise) trading. More precise pricing can result in reduced liquidity trading, and in an extreme case, in market failure, where informed traders are unwilling to share risk with others by trading. There may thus be an optimal level of transparency.

According to Rindi (2008), given the amount of informed traders, a more transparent market is more liquid. This is in line with most earlier theoretical work. However, with endogenous information acquisition, the analysis is more complex. According to Rindi (2008), limit orders accommodate liquidity shocks caused by liquidity traders. Informed limit order traders are of course in a better position to do this, since they do not face adverse selection costs. Uninformed traders are reluctant to accommodate large market orders, in fear of facing an informed trader. Greater transparency reduces the incentive to acquire information, and thus diminishes the number of liquidity traders.

In order to test the effects of changes in market transparency for informed trading, we estimate the Probability of Informed Trading (PIN) model of Equation 2 above. We perform the estimation for both the pre and post-change periods.

Table 5 presents the results[10]. For most stocks, the estimated PIN value falls between approximately 10% and 20%. There is no clear change with the advent of anonymous trading, contrary to our expectations. The average probability of informed trading coefficient is slightly lower after the change, but not significantly so. Also, there are as many stocks for which PIN increases, as there are stocks with a decrease in PIN.

E. Informed trading as an explanation for changes in bid-ask spreads

A testable prediction of the Foucault, Moinas, and Theissen (2007) model is that in the presence of few informed traders (a low participation rate by informed traders), a switch to anonymity from a transparent limit order book results in lower bid-ask spreads. If the number of informed traders exceeds a certain limit, the effect is reversed, i.e. bid-ask spreads will increase.

This regression analysis is similar in many ways to the analysis of Section IV.B above. The dependent variable is the bid-ask spread, measured both in euros and as a percentage. The explanatory variables are the same as above, with the addition of the Probability of Informed Trading, PIN. In other words, our explanatory variables are the logarithmic trading volume, the average trade price, 30-minute return volatility, a pre-post change dummy, and the PIN[11].

---

[10] The maximum likelihood model converges for all stocks and both sample periods, with one exception, Nokian Renkaat Oyj, NRE1V, in the post-change period.

[11] Of course, this is only an approximate test of the Foucault, Moinas, and Theissen (2007) hypothesis. Their conclusions apply with certain values of the parameter β for the participation rate of informed

This regression is not directly comparable with the analysis presented in Section 4.B above, however. Instead of a pooled regression with values for all dates and all stocks in the sample, this is a cross-sectional regression over all stocks. We calculate the average value of each variable, separately for the pre- and post-change periods (transparent and non-transparent trading systems, respectively). This results in two estimated values for each stock. The reason for using this smaller number of observations is the impossibility of obtaining daily estimates of the Probability of Informed Trading. As we pointed out in Section II above, a minimum of 60 trading days is generally regarded as necessary for a reliable estimation of the PIN measure. We therefore have two values for the PIN for each stock, one for the pre-change period, and another for the post-change period.

We estimate the following model:

$$s_{i,t} = \gamma_0 + \gamma_1 \log\left(V_{i,t}\right) + \gamma_2 P_{i,t} + \gamma_3 \sigma_{i,t} + \gamma_4 D_t^{post} + \gamma_5 PIN_t + \varepsilon_{i,t}\,, \tag{7}$$

where *s* is the spread of stock *i* at time *t*, *V* is the trading volume of the stock, *P* the average stock price, σ is midquote return volatility; D^{post} is a dummy variable that takes on a value of 0 for the period before the changes (the transparent market), and a value of 1 after the changes (the opaque market); *PIN* is the probability of informed trading. There are only time time periods time *t*: the pre-change and the post-change time periods.

Table 6 presents the results of the regression of Equation 7. Trading volume and stock price are significant explanatory variables in all regressions. Both also have the expected sign, similarly to the earlier analysis of Section 4.B: the greater the trading volume and the greater the price, the smaller is the bid-ask spread. For the regression in Panel B, where the bid-ask spread is measured in euros, the stock price has the opposite sign, which is also expected. The PIN variable is not significant in any of these regressions, albeit it does have the expected

---

traders. In other words, when $\beta \leq \beta^{**}$, expected bid-ask spreads are reduced. If $\beta > \beta^{*}$, expected bid-ask spreads are larger after a switch to anonymity. In between these cases, when $\beta^{**} < \beta \leq \beta^{*}$, small trade spreads are reduced, while large trade spreads are increased.

sign, being positive. However, the last column is a regression with the PIN variable and the dummy variable only. Here the PIN is both significant, and has the expected sign, being positive. However, since this regression does not control for any of the other variables, we do not feel that this result is valid.

As a robustness test, we also use a smaller sample of the largest and most liquid stocks. We select the 20 most liquid companies, excluding Nokia, which is in a class of its own, measured in trading volume and market capitalization. We run the same regression model as above. The results are substantially the same, and are available upon request.

V. Conclusions

In this paper, we study a change from a fully pre-trade transparent limit order book to a less transparent one, at the OMX Helsinki Stock Exchange (the former HEX). The exchange changed their trading system in March 2006, by eliminating the display of buyer and seller identities in the electronic limit order book. However, the market remained fully post-trade transparent, in that the identities of the buying and selling counterparties are known to all market participants immediately after a trade takes place.

Our main results are the following. We test a theoretical prediction of Foucault, Moinas, and Theissen (2007), using Probability of Informed Trading (PIN) estimates as a proxy for the participation rate of informed traders. We use several trading related variables, such as trading volume, stock price, intraday return volatility, the estimated values of PIN, and a period dummy variable. In a regression with stock-specific averages for these explanatory variables, for the pre-change and post-change periods, we do not find evidence of any significant explanatory power for the PIN variable.

We also study the broader question of the effects of a change from a pre-trade transparent to an anonymous electronic limit order book market. The unresolved question, posed by many market regulators and exchange officials the world over, is whether transparency improves market quality. It is commonly agreed that market quality improves with a decrease in the bid-ask spread, an increase in trading volume, and a decrease in return volatility.

We find that the effects of a switch to a less pre-trade transparent limit order book are not unambiguously positive. Our analysis shows no significant change in average bid-ask spreads, and a significant increase in intraday volatility. However, trading volume increases after the changes, mitigating the two negative effects. In a cross-sectional analysis, using trading volume, stock price, and intraday volatility as explanatory variables for the bid-ask spread, we conclude that the change itself is a significant factor in explaining changes in market quality after the switch.

As a comparison with the changes in transparency in the electronic limit order book market, we also study trading in the upstairs market. The changes in the trading system do not affect this segment, where pre-trade anonymity does not exist. We find that there indeed is no change in the number of block trades and contract trades. Neither does the internalization rate of trades change.

VI. References

Admati, Anat R., and Paul Pfleiderer, 1991, Sunshine Trading and Financial Market Equilibrium, *Review of Financial Studies* 4, 443-481.

Baruch, Shmuel, 2005, Who Benefits from an Open Limit-Order Book?, *Journal of Business* 78, 1267-1306.

Bessembinder, Hendrik, 2003, Issues in assessing trade execution costs, *Journal of Financial Markets* 6, 233–257.

Bessembinder, Hendrik, and Kumar Venkataraman, 2004, Does an electronic stock exchange need an upstairs market?, *Journal of Financial Economics* 73, 3-36.

Bloomfield, Robert, and Maureen O'Hara, 1999, Market transparency: Who wins and who loses?, *Review of Financial Studies* 12, 5-35.

Bloomfield, Robert, and Maureen O'Hara, 2000, Can transparent markets survive?, *Journal of Financial Economics* 55, 425-459.

Bloomfield, Robert, Maureen O'Hara, and Gideon Saar, 2005, The 'Make' or 'Take' Decision in an Electronic Market: Evidence on the Evolution of Liquidity, *Journal of Financial Economics* 75, 165–200.

Boehmer, Ekkehart, Joachim Grammig, and Erik Theissen, 2007, Estimating the probability of informed trading – does trade misclassification matter?, *Journal of Financial Markets* 10, 26–47.

Boehmer, Ekkehart, Gideon Saar, and Lei Yu, 2005, Lifting the Veil: An Analysis of Pre-trade Transparency at the NYSE, *Journal of Finance* 60, 783-815.

Booth, G. Geoffrey, Juha-Pekka Kallunki, Ji-Chai Lin, and Teppo Martikainen, Internalization and stock price clustering: Finnish evidence, *Journal of International Money and Finance* 19, 737–751.

Chakrabarty, Bidisha, Bingguang Li, Vanthuan Nguyen, and Robert A. Van Ness, 2007, Trade classification algorithms for electronic communications network trades, *Journal of Banking and Finance* 31, 3806-3821.

Chan, Louis K.C., and Josef Lakonishok, 1995, The behavior of stock prices around institutional trades, *Journal of Finance* 50, 1147-1174.

Chowdhry, Bhagwan, and Vikram Nanda, 1991, Multimarket trading and market liquidity, *Review of Financial Studies* 4, 483-511.

Comerton-Forde, Carole, Alex Frino, and Vito Mollica, 2005, The impact of limit order anonymity on liquidity: Evidence from Paris, Tokyo and Korea, *Journal of Economics and Business* 57, 528-540.

Copeland, Thomas E., and Dan Galai, 1983, Information Effects and the Bid-Ask Spreads, *Journal of Finance* 38, 1457–69.

Easley, David, Sören Hvidkjaer, and Maureen O’Hara, 2002, Is information risk a determinant of asset returns?, *Journal of Finance* 57, 2185-2221.

Easley, David, Sören Hvidkjaer, and Maureen O’Hara, 2005, Factoring information into returns, (July 27, 2005), EFA 2004 Maastricht Meetings Paper No. 4118, http://ssrn.com/abstract=556079.

Easley, David, Nicholas M. Kiefer, Maureen O'Hara and Joseph B. Paperman, 1996. Liquidity, Information, and Infrequently Traded Stocks, *Journal of Finance* 51, 1405-1436.

Easley, David, Nicholas M. Kiefer, and Maureen O'Hara, 1996, Cream-Skimming or Profit-Sharing? The Curious Role of Purchased Order Flow, *Journal of Finance* 51, 811-833.

Easley, David, Nicholas M. Kiefer, and Maureen O'Hara, 1997, One Day in the Life of a very Common Stock, *Review of Financial Studies* 10, 805-835.

Easley, David and Maureen O’Hara, 1992, Time and the Process of Security Price Adjustment, *Journal of Finance* 57, 577–605.

Easley, David, Maureen O'Hara, and Joseph Paperman, 1998, Financial analysts and information-based trade, *Journal of Financial Markets* 1, 175-201.

Easley, David, Maureen O'Hara, and Gideon Saar, 2001, How Stock Splits Affect Trading: A Microstructure Approach, *Journal of Financial and Quantitative Analysis* 36, 25-51.

Ellis, Katrina, Roni Michaely, and Maureen O'Hara, 2000, The accuracy of trade classification rules: Evidence from Nasdaq, *Journal of Financial and Quantitative Analysis* 35, 529-551.

Finucane, Thomas J., 2000, A direct test of methods for inferring trade direction from intra-day data, *Journal of Financial and Quantitative Analysis* 35, 553-576.

Flood, Mark D., Ronald Huisman, Kees G. Koedijk, Ronald J. Mahieu, 1999, Quote disclosure and price discovery in multiple-dealer financial markets, *Review of Financial Studies* 12, 37-59.

Foucault, Thierry, Sophie Moinas, and Erik Theissen, 2007, Does anonymity matter in electronic limit order markets?, *Review of Financial Studies* 20, 1707-1747.

Garfinkel, Jon A., and M. Nimalendran, 2003, Market Structure and Trader Anonymity: An Analysis of Insider Trading, *Journal of Financial and Quantitative Analysis* 8, 591–610.

Glosten, Lawrence R., and Paul R. Milgrom, 1985, *Journal of Financial Economics* 14, 71-100.

Grammig, Joachim, Dirk Schiereck, Erik Theissen, 2001, Knowing me, knowing you: Trader anonymity and informed trading in parallel markets, *Journal of Financial Markets* 4, 385-412.

Harris, Jeffrey H. and Paul H. Schultz, 1997, The importance of firm quotes and rapid executions: Evidence from the January 1994 SOES rules change, *Journal of Financial Economics* 45, 135-166.

Hendershott, Terrence, and Charles M. Jones, 2005, Island goes dark: Transparency, fragmentation, and regulation, *Review of Financial Studies* 18, 743-793.

Huang, Roger D. and Hans R. Stoll, 1996, Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE, *Journal of Financial Economics* 41, 313-357.

de Jong, Frank, Theo Nijman and Ailsa Roell, 1996, Price Effects of Trading and Components of the Bid-ask Spread on the Paris Bourse, *Journal of Empirical Finance* 3, 193-213.

Jones, Charles M., Gautam Kaul, and Marc L. Lipson, 1994, Transactions, Volume and Volatility, *Review of Financial Studies* 7, 631–51.

Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1336.

Lee, Charles M. C., and Mark J. Ready, 1991, Inferring Trade Direction from Intraday Data, *Journal of Finance* 46, 733-746.

Madhavan, Ananth, 1995, Consolidation, Fragmentation, and the Disclosure of Trading Information, *Review of Financial Studies* 8, 579-603.

Madhavan, Ananth, 1996, Security Prices and Market Transparency, *Journal of Financial Intermediation* 5, 255-283.

Madhavan, Ananth, and Minder Cheng, 1997, In Search of Liquidity: Block Trades in the Upstairs and Downstairs Markets, *Review of Financial Studies* 10,175–203.

Madhavan, Ananth, David Porter, and Daniel Weaver, 2005, Should securities markets be transparent?, *Journal of Financial Markets* 8, 266–288.

Pagano, Marco and Ailsa Roell, 1996, Transparency and Liquidity: A Comparison of Auction and Dealer Markets with Informed Trading, *Journal of Finance* 51, 579-611

Rindi, Barbara, 2008, Informed Traders as Liquidity Providers: Anonymity, Liquidity and Price Formation, *Review of Finance* 12, 497-532.

Savickas, Robert, and Arthur J. Wilson, 2003, On Inferring the Direction of Option Trades, *Journal of Financial and Quantitative Analysis* 38, 881-902.

Seppi, Duane J., 1990, Equilibrium Block Trading and Asymmetric Information, *Journal of Finance* 45, 73–94.

Simaan, Yusif, Daniel G. Weaver, and David K. Whitcomb, 2003, Market maker quotation behavior and pretrade transparency, *Journal of Finance* 58, 1247-1267.

Smith, Brian F., Alasdair S. Turnbull, and Robert W. White, 2001, Upstairs Market for Principal and Agency Trades: Analysis of Adverse Information and Price Effects, *Journal of Finance* 56, 1723-1746.

Theissen, Erik, 2003, Trader Anonymity, Price Formation and Liquidity, *European Finance Review* 7, 1–26.

VII. Tables

Table 1

Descriptive statistics, stock sample

This table presents descriptive statistics for our sample of stocks. The sample period is December 13, 2005 – May 13, 2006. *Ticker code* is the official stock ID in the trading system, *Market cap* is the market capitalization as of end of 2005, *Number of trades* is the average daily number of trades, *Price* is the average price over the entire sample period, and *Trading volume* is the average daily trading volume, in millions of euros, *Trade size* is the average number of shares per trade.

For each descriptive variable, we also report average, median, standard deviation, minimum, and maximum values, the skewness coefficient, and excess kurtosis.

Stock name	Ticker code	Market cap	Number of trades	Price	Trading volume	Trade size
Alma Media Oyj	ALN1V	878	56	7.66	1.25	3423.3
Amer Sports Corporation	AMEAS	1,124	321	16.48	3.65	690.0
Cargotec Oyj	CGCBV	1,866	388	33.59	5.66	447.0
Elcoteq SE A	ELQAV	577	201	18.45	2.14	493.4
Elisa Oyj	ELI1V	2,596	848	16.35	14.16	1012.5
F-Secure Oyj	FSC1V	455	167	2.70	1.24	2376.3
Finnair Oyj	FIA1S	643	135	12.56	1.76	887.6
Finnlines Oyj	FLG1S	543	59	14.99	1.65	1880.4
Fortum Oyj	FUM1V	13,864	1,638	19.17	52.09	1610.6
Huhtamäki Oyj	HUH1V	1,374	309	14.85	3.64	814.4
KONE Oyj	KNEBV	4,261	653	34.05	9.65	426.4
Kemira GrowHow Oyj	KGH1V	899	195	5.52	1.40	1103.9
Kemira Oyj	KRA1V	797	270	13.88	3.50	919.2
Kesko Oyj B	KESBV	2,310	512	26.23	5.76	410.9
M-real Oyj B	MRLBV	1,384	536	4.57	8.41	3162.0
Metso Oyj	MEO1V	3,274	1,124	28.62	24.44	728.2
Neste Oil Oyj	NES1V	6,122	1,252	26.22	29.99	902.7
Nokia Oyj	NOK1V	64,463	6,345	16.35	515.47	4711.7
Nokian Renkaat Oyj	NRE1V	1,289	912	13.38	12.43	969.9
Nordea Bank AB (publ) FDR	NDA1V	22,729	346	9.45	13.34	3952.3
OKO Pankki Oyj	OKOAS	2,385	406	12.85	5.25	1018.4
Outokumpu Oyj	OUT1V	2,272	830	15.85	14.82	1134.1
Perlos Oyj	POS1V	565	355	7.67	2.88	972.0
Ramirent Oyj	RMR1V	378	115	26.73	1.91	669.3
Rautaruukki Oyj K	RTRKS	2,801	1,035	26.09	16.32	573.3
Sampo Oyj A	SAMAS	8,307	1,346	16.23	39.73	1737.0
Sponda Oyj	SDA1V	344	96	8.44	1.15	1405.7
Stockmann Oyj Abp B	STCBV	880	202	32.77	2.37	368.7

Stora Enso Oyj R	STERV	9,021	1,368	11.80	45.15	2573.7
TeliaSonera AB	TLS1V	20,364	316	4.72	8.02	5254.9
TietoEnator Oyj	TIE1V	1,847	870	28.77	16.63	604.5
UPM-Kymmene Oyj	UPM1V	8,662	1,675	17.76	60.66	1849.7
Uponor Oyj	UNR1V	988	204	21.37	2.44	533.1
Wärtsilä Oyj Abp	WRTBV	2,353	599	29.57	9.40	519.7
YIT Oyj	YTY1V	2,254	635	31.83	10.13	540.5
Mean		5568	752	17.93	27.10	1447.9
Median		1866	406	16.35	8.02	969.9
Standard deviation		11564	1077	9.16	86.33	1261.2
min		344	56	2.70	1.15	368.7
max		64463	6345	34.05	515.47	5254.9
skewness		4.26	4.35	0.24	5.64	1.70
kurtosis		20.62	22.32	-1.00	32.66	2.26

Table 2

Descriptive statistics, pre-change (non-anonymous) and post-change (anonymous) periods

This table presents descriptive statistics for pre and post-change periods (transparent and anonymous trading systems, respectively). Bid-ask spreads (in euros and in percent) are time-weighted bid-ask spreads during continuous trading hours. Trading volume is average daily trading volume in thousands of euros, and in thousands of shares. Volatility is the standard deviation of asset returns on a 30-minute time interval. We report the values calculated for two periods: "pre" and "post", where the former refers to the period of transparent trading before the changes made in March 2006, and the latter to the period of anonymous trading after the changes. We have two sample periods of 62 days each.

We report test statistics (respectively the Wilcoxon z test and the paired Student t test) for the null hypothesis of equal median and mean values, respectively, for the pre and post periods. *, **, and *** indicate significance on a 10%, 5%, and 1% level, respectively. P-values are in parentheses.

	Mean			Median		
	Pre	Post	t-test	Pre	Post	Wilcoxon
Bid-ask spread, euros	0.0410	0.0408	0.545	0.0358	0.0348	0.0475**
Bid-ask spread, percent	0.284%	0.281%	0.45136	0.214%	0.207%	0.0475**
Number of trades	673.14	831.29	0.0005***	423.58	467.16	0.0002***
Trade price	17.65	18.19	0.3347	16.03	16.65	0.0168**
Volume, Number of shares, 1000s	1,596.80	1,776.00	0.0787	338.57	443.48	0.0259**
Volume, euros, 1000s	24,390	29,808	0.0860	6,803	9,177	0.0037***
Trade size	1,531	1,367	0.0196**	1,003	897	0.0054***
Volatility	0.322%	0.367%	2.59E-05***	0.282%	0.325%	4.53E-05***

N=35.

Table 3

Multivariate regression results

This table presents the coefficient estimates of the regression described in Section IV.B. The regression model is

$$s_{i,t} = \gamma_0 + \gamma_1 \log\left(V_{i,t}\right) + \gamma_2 P_{i,t} + \gamma_3 \sigma_{i,t} + \gamma_4 D_t^{post} + \varepsilon_{i,t}.$$

The dependent variable is the spread, measured either in euros in or as a percentage. The regression is a pooled regression over all stocks. The explanatory variables are *log(V),* the logarithm of average trading volume in euros, *P* is the average price, σ is the standard deviation of 30-minute logarithmic returns, and D^{post} takes the value 0 for the pre-change (transparent) market, and the the value 1 for the post-change (anonymous) market. All variables are calculated per stock and per trading day. The fixed effects regressions include stock specific dummy variables (omitted in this table). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Figures in parenthesis are t-statistics. The number of observations is 92.

	Pooled regression		Pooled regression, fixed effects	
	Percentage spread	Spread in euros	Percentage spread	Spread in euros
Constant	0.008***	0.085***	0.003***	0.017***
	(56.66)	(32.89)	(13.98)	(3.14)
Volatility	0.091***	1.506***	0.064***	1.436***
	(13.73)	(12.49)	(10.62)	(9.98)
Log(volume)	-0.000***	-0.005***	-0.000***	-0.002***
	(-43.08)	(-34.25)	(-8.24)	(-8.19)
Price	-0.000***	0.002***	-0.000	0.002***
	(-22.06)	(65.92)	(-0.68)	(13.50)
Pre/post dummy	0.000***	0.001**	0.000*	-0.000
	(3.813)	(2.054)	(1.759)	(-0.256)
Adj. R^2	0.550	0.780	0.854	0.877

Table 4

The upstairs market and internalization rates

This table presents internalization rates for the main non-continuous trade types, block trades and contract trades. Internal trades have the same buyer and seller. Internalization rate is the percentage of internal trades out of all trades. The numbers for block and contract trades are total numbers of trades in the pre- and post –change periods.

The t-test is for the equality of the mean for the internalization rate of the pre-change period and the post-change period. We do no reject the null hypothesis of equal means.

	Pre change (transparent market)						Post change (anonymous market)					
	Block trades		Contract trades		Internalization rate		Block trades		Contract trades		Internalization	
Stock	internal	external	internal	external	Block trades	Contract trades	internal	external	internal	external	Block trades	Contract trades
Amer Sports Oyj	12	0	272	3	100.0%	98.9%	2	0	114	2	100.0%	98.3%
Cargotec Oyj	6	0	163	0	100.0%	100.0%	13	0	214	1	100.0%	99.5%
Elisa Oyj	26	1	425	5	96.3%	98.8%	13	0	400	3	100.0%	99.3%
Fortum Oyj	81	0	730	3	100.0%	99.6%	139	1	1647	3	99.3%	99.8%
Huhtamäki Oyj	2	0	171	1	100.0%	99.4%	6	1	148	3	85.7%	98.0%
Kesko Oyj B	5	0	300	1	100.0%	99.7%	11	0	180	3	100.0%	98.4%
Kone Oyj B	14	0	182	2	100.0%	98.9%	27	0	300	0	100.0%	100.0%
Metso Oyj	47	0	404	0	100.0%	100.0%	35	0	496	7	100.0%	98.6%
M-real Oyj B	5	0	199	1	100.0%	99.5%	34	0	303	0	100.0%	100.0%
Neste Oil Oyj	66	0	517	1	100.0%	99.8%	24	1	410	0	96.0%	100.0%
Nokia Oyj	541	16	2953	20	97.1%	99.3%	513	17	2663	9	96.8%	99.7%
Nokian Renkaat Oyj	26	1	705	5	96.3%	99.3%	12	4	380	9	75.0%	97.7%
Nordea Bank AB (publ) FDR	7	0	139	9	100.0%	93.9%	20	0	123	1	100.0%	99.2%
OKO Bank Oyj A	11	0	280	4	100.0%	98.6%	2	0	97	1	100.0%	99.0%
Outokumpu Oyj	24	0	441	7	100.0%	98.4%	22	0	437	5	100.0%	98.9%

Rautaruukki Oyj K	24	1	414	1	96.0%	99.8%	24	2	266	3	92.3%	98.9%
Sampo Oyj A	74	1	781	1	98.7%	99.9%	73	9	699	4	89.0%	99.4%
Stora Enso Oyj R	84	2	627	2	97.7%	99.7%	99	1	587	0	99.0%	100.0%
TeliaSonera AB	11	0	127	1	100.0%	99.2%	20	1	110	0	95.2%	100.0%
Tietoenator Oyj	35	1	341	2	97.2%	99.4%	30	1	423	4	96.8%	99.1%
UPM-Kymmene Oyj	124	2	575	7	98.4%	98.8%	116	5	590	1	95.9%	99.8%
Wärtsilä Oyj Abp B	17	0	322	0	100.0%	100.0%	12	0	284	2	100.0%	99.3%
YIT-Yhtymä Oyj	15	1	241	5	93.8%	98.0%	15	0	497	1	100.0%	99.8%
Average	54.7	1.1	491.7	3.5	98.8%	99.1%	54.9	1.9	494.3	2.7	96.6%	99.2%
t test											1.796	0.531
											(0.086)	(0.601)

Table 5

Probability of Informed Trading

This table presents the results of the PIN (Probability of Informed Trading) analysis. PIN is defined as

$$PIN = \frac{\alpha\mu}{\alpha\mu + \varepsilon_s + \varepsilon_b},$$

where α is the probability of an information event, μ is the Poisson parameter for the arrival of informed trades, and ε_s and ε_b are the Poisson densities for the arrival of uninformed sell and buy trades, respectively. *Std error* (in parenthesis) are the standard error statistics of the PIN estimates. In the table below, δ is the probability of negative news. There is only one value for the ε variable since we make the assumption that uninformed traders are equally likely to be buyers and sellers.

The maximum likelihood model does not converge in a few cases. These values are excluded from our analysis of changes in PIN. The last column indicates whether the estimated PIN has increased or decreased from the first period. We report the average, median, minimum, and maximum values, as well as standard deviation statistics (in parenthesis) for all variables.

The last row presents a paired two-sided t-test for the PIN estimates of the pre-change period and the post-change period. The value in parenthesis is the t-probability value of the t-test.

	Pre					Post					
Stock ticker	Alpha	delta	mu	epsilon	PIN	alpha	delta	mu	epsilon	PIN	Change in PIN
ALN1V	0.327	0.867	45.48	27.15	0.215	0.486	0.333	20.60	20.33	0.198	-
	(0.062)	(0.080)	(2.23)	(0.55)	(0.032)	(0.089)	(0.097)	(1.59)	(0.62)	(0.029)	
AMEAS	0.434	0.485	136.53	142.33	0.172	0.450	0.212	142.96	136.55	0.191	+
	(0.065)	(0.098)	(4.03)	(1.36)	(0.021)	(0.063)	(0.078)	(3.45)	(1.22)	(0.022)	
CGCBV	0.145	0.778	242.56	147.59	0.107	0.387	0.458	206.17	204.62	0.163	+
	(0.045)	(0.139)	(6.76)	(1.13)	(0.029)	(0.062)	(0.102)	(4.38)	(1.43)	(0.022)	
ELQAV	0.194	0.586	217.23	84.04	0.201	0.208	0.500	206.05	92.35	0.188	-
	(0.061)	(0.191)	(6.66)	(0.85)	(0.051)	(0.052)	(0.141)	(6.51)	(0.99)	(0.038)	
ELI1V	0.501	0.386	254.28	369.70	0.147	0.564	0.398	244.20	377.75	0.154	+

	(0.064)	(0.088)	(4.95)	(2.01)	(0.016)	(0.063)	(0.083)	(4.71)	(2.07)	(0.015)	
FSC1V	0.212	0.718	126.67	77.24	0.148	0.250	0.585	121.02	65.48	0.188	+
	(0.048)	(0.109)	(4.87)	(1.00)	(0.028)	(0.060)	(0.143)	(4.14)	(0.77)	(0.037)	
FIA1S	0.422	0.343	82.60	60.41	0.224	0.403	0.360	86.03	46.27	0.273	+
	(0.063)	(0.094)	(2.57)	(0.80)	(0.027)	(0.062)	(0.096)	(2.41)	(0.69)	(0.031)	
FLG1S	0.369	0.557	32.63	19.40	0.237	0.444	0.277	42.63	27.78	0.254	+
	(0.066)	(0.108)	(1.85)	(0.48)	(0.033)	(0.067)	(0.087)	(1.92)	(0.60)	(0.029)	
FUM1V	0.569	0.419	382.15	639.39	0.145	1.000	0.002	213.07	246.62	0.302	+
	(0.064)	(0.085)	(6.42)	(2.94)	(0.014)	.	.	.	.	.	
HUH1V	0.441	0.595	100.77	111.39	0.166	0.440	0.184	140.40	159.55	0.162	-
	(0.064)	(0.097)	(3.13)	(1.11)	(0.021)	(0.064)	(0.074)	(3.84)	(1.36)	(0.020)	
KNEBV	0.416	0.457	181.63	259.91	0.127	0.306	0.527	258.79	344.04	0.103	-
	(0.063)	(0.099)	(4.66)	(1.70)	(0.017)	(0.059)	(0.115)	(5.98)	(1.82)	(0.018)	
KGH1V	0.315	0.564	100.38	91.61	0.147	0.500	0.565	108.66	74.34	0.268	+
	(0.066)	(0.117)	(5.36)	(1.23)	(0.025)	(0.064)	(0.134)	(4.12	(0.84)	(0.026)	
KRA1V	0.211	0.125	168.07	114.21	0.135	0.464	0.429	115.91	122.08	0.180	+
	(0.052)	(0.044)	(5.48)	(1.05)	(0.029)	(0.067)	(0.093)	(3.75)	(1.34)	(0.021)	
KESBV	0.607	0.709	157.39	177.38	0.212	0.372	0.482	247.22	268.66	0.146	-
	(0.063)	(0.074)	(3.41)	(1.52)	(0.018)	(0.062)	(0.105)	(5.22)	(1.66)	(0.021)	
MRLBV	0.119	0.000	242.92	168.37	0.079	0.258	0.188	296.95	279.24	0.121	+
	(0.042)	.	(7.76)	(1.23)	(0.026)	(0.056)	(0.098)	(6.34)	(1.62)	(0.023)	
MEO1V	0.274	0.823	436.15	415.85	0.126	0.419	0.308	414.45	590.20	0.128	+
	(0.057)	(0.093)	(7.39)	(1.97)	(0.023)	(0.063)	(0.091)	(6.74)	(2.47)	(0.017)	
NES1V	0.424	0.600	378.26	482.69	0.142	0.274	0.710	470.48	577.71	0.100	-
	(0.064)	(0.098)	(6.34)	(2.29)	(0.019)	(0.054)	(0.077)	(7.63)	(2.36)	(0.018)	
NOK1V	0.431	0.012	1453.06	2545.42	0.110	0.309	0.021	1672.43	2904.86	0.082	-
	(0.065)	(0.009)	(13.71)	(5.29)	(0.015)	(0.062)	(0.015)	(17.33)	(5.59)	(0.015)	
NRE1V	0.328	0.251	326.09	425.27	0.112						-

	(0.060)	(0.097)	(6.59)	(2.06)	(0.018)						
NDA1V	0.210	0.500	182.30	133.22	0.125	0.551	0.119	155.77	150.91	0.221	+
	(0.052)	(0.139)	(5.27)	(1.11)	(0.027)	(0.064)	(0.056)	(3.46)	(1.39)	(0.021)	
OKOAS	0.306	0.684	230.40	196.90	0.152	0.353	0.518	148.63	163.94	0.138	-
	(0.059)	(0.107)	(4.99)	(1.38)	(0.025)	(0.061)	(0.110)	(4.18)	(1.30)	(0.021)	
OUT1V	0.377	0.435	260.07	274.19	0.152	0.355	0.409	391.03	451.45	0.133	-
	(0.062)	(0.103)	(5.10)	(1.66)	(0.021)	(0.061)	(0.105)	(6.65)	(2.12)	(0.020)	
POS1V	0.413	0.447	143.58	131.47	0.184	0.295	0.278	179.73	122.90	0.178	-
	(0.065)	(0.100)	(3.92)	(1.26)	(0.024)	(0.058)	(0.106)	(4.29)	(1.09)	(0.029)	
RAIVV	0.145	0.332	127.37	47.78	0.162	0.496	0.125	48.00	38.97	0.234	+
	(0.045)	(0.158)	(4.49)	(0.65)	(0.042)	(0.067)	(0.046)	(2.01)	(0.72)	(0.025)	
RMR1V	0.334	0.654	61.66	40.48	0.203	0.350	0.461	82.20	55.74	0.205	+
	(0.062)	(0.107)	(2.77)	(0.69)	(0.030)	(0.061)	(0.108)	(2.83)	(0.78)	(0.029)	
RTRKS	0.371	0.477	269.73	376.12	0.118	0.492	0.200	343.61	541.91	0.135	+
	(0.061)	(0.105)	(5.70)	(1.94)	(0.017)	(0.064)	(0.073)	(5.96)	(2.43)	(0.015)	
SAMAS	0.339	0.237	410.97	524.32	0.117	0.516	0.313	415.55	684.45	0.136	+
	(0.060)	(0.082)	(7.07)	(2.26)	(0.018)	(0.063)	(0.082)	(6.47)	(2.73)	(0.015)	
SDA1V	0.505	0.565	44.10	35.67	0.238	0.509	0.500	46.91	45.30	0.208	-
	(0.067)	(0.092)	(1.87)	(0.69)	(0.025)	(0.067)	(0.092)	(2.00)	(0.77)	(0.023)	
STCBV	0.455	0.649	84.34	82.65	0.189	0.502	0.389	78.07	90.23	0.179	-
	(0.065)	(0.090)	(2.83)	(0.99)	(0.022)	(0.077)	(0.092)	(3.68)	(1.43)	(0.022)	
STERV	0.356	0.009	443.80	473.81	0.143	0.412	0.404	389.63	697.19	0.103	-
	(0.062)	(0.004)	(7.12)	(2.23)	(0.022)	(0.063)	(0.099)	(7.52)	(2.77)	(0.014)	
TLS1V	0.695	0.063	109.37	110.21	0.257	0.258	0.250	175.30	156.33	0.127	-
	(0.063)	(0.039)	(2.84)	(1.42)	(0.018)	(0.056)	(0.153)	(4.80)	(1.21)	(0.024)	
TIE1V	0.484	0.338	257.19	111.24	0.359	0.450	0.305	327.39	513.13	0.126	-
	(0.063)	(0.080)	(3.70)	(1.09)	(0.031)	(0.070)	(0.096)	(6.72)	(2.60)	(0.017)	
UPM1V	0.271	0.001	500.18	619.45	0.099	0.532	0.075	451.20	839.44	0.125	+

	(0.058)	(0.001)	(8.78)	(2.47)	(0.019)	(0.063)	(0.030)	(6.95)	(3.04)	(0.013)	
UNR1V	0.250	0.047	72.68	70.84	0.114	0.025	0.675	166.50	109.64	0.019	-
	(0.049)	(0.017)	(2.76)	(0.86)	(0.020)	(0.006)	(0.113)	(4.81)	(1.10)	(0.004)	
WRTBV	0.507	0.692	197.60	217.48	0.187	0.451	0.250	223.66	297.55	0.145	-
	(0.064)	(0.085)	(4.04)	(1.58)	(0.020)	(0.063)	(0.079)	(4.74)	(1.79)	(0.018)	
YTY1V	0.499	0.711	164.18	199.23	0.171	0.373	0.393	271.71	373.84	0.119	-
	(0.064)	(0.082)	(3.75)	(1.48)	(0.018)	(0.062)	(0.103)	(6.20)	(2.02)	(0.018)	
Mean					0.166					0.162	
Median					0.150					0.150	
Minimum					0.079					0.019	
Maximum					0.359					0.302	
Standard deviation					0.056					0.059	
t-test										0.481	
										(0.713)	

Table 6

Test of the Foucault, Moinas, and Theissen (2007) model using the pre-change PIN

This table presents the results of our analysis of informed trading as an explanatory variable of market quality. *Percentage spread* is the average percentage quoted bid-ask spread, and the *Spread in euros* is the average difference between the bid and the ask prices.

The explanatory variables are *Log volume*, the logarithm of average trading volume in euros, *Price* is the average price, *Volatility* is the standard deviation of 30-minute logarithmic returns, and *PIN* is the Probability of Informed Trading. All variables are stock-specific averages for the pre-change and the post-change periods. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Percentage spread as the dependent variable

	(1)	(2)	(3)	(4)	(5)
Constant	0.016***	0.018***	0.018***	0.017***	0.0007
	9.321	13.256	13.113	9.869	1.029
Log volume	-0.001***	-0.001***	-0.001***	-0.001***	
	-8.921	-10.461	-10.386	-9.037	
Price	-4.5E-05***	-4.0E-05***	-4.0E-05***	-3.0E-05***	
	-3.277	-2.931	-2.914	-2.904	
Volatility	0.181*				
	1.752				
PIN	0.003			0.003	0.013***
	1.101			1.253	3.820
Dummy	0.000		0.000		3.4E-05
	-0.086		0.216		0.086
R^2	0.681	0.672	0.667	0.675	0.158

Panel B: Spread in euros as the dependent variable

	(1)	(2)	(3)	(4)	(5)
Constant	0.198***	0.210***	0.211***	0.191***	0.019*
	6.544	9.015	8.936	6.543	1.808
Log volume	-0.012***	-0.013***	-0.013***	-0.012***	
	-7.274	-8.482	-8.415	-7.271	
Price	0.002***	0.002***	0.002***	0.002***	
	8.080	8.108	8.048	8.150	
Volatility	-1.743				
	-0.988				
PIN	0.048			0.044	0.135**
	1.174			1.098	2.30
Dummy	0.000		-0.001		0.0003
	0.085		-0.146		0.047
R^2	0.617	0.622	0.616	0.623	0.049

N=68.

VIII. Figures

Figure 1 Tree diagram of the trading process, from Easley, Kiefer, and O'Hara (1997).

This diagram is a representation of the trading model discussed in Section III.A. α is the probability of an information event, δ is the probability of a low signal, μ is the probability that the trade comes from an informed trader, ½ is the probability that an uninformed trader is a seller, and ε is the probability that the uninformed trader will actually trade. Nodes to the left of the dotted line occur only at the beginning of the trading day; nodes to the right are possible at each trading interval.