Liquidity and Information-Based Trading on the Order Driven Capital Market: The Case of the Prague Stock Exchange

Libor Němeček

Center for Economic Research and Graduate Education, Charles University
and
The Economic Institute of the Academy of Sciences of the Czech Republic
Politickych veznu 7
Prague 1, 111 21
Czech Republic
fax: (+422) 24 22 7143
e-mail: libor.nemecek@cerge.cuni.cz

First Draft: March 3, 1997
This Version: August 25, 1997 (CERGE-EI WP)

Abstract: This paper investigates the relation between liquidity and information-based trading in the context of an order-driven auction. A model similar in spirit to that of Easley et al. (1996) is used to determine how often new information occurs and how it influences the composition of orders submitted to the market. The risk of information-based trading is estimated for a sample of Prague Stock Exchange listed stocks. The empirical results agree with previous findings that the risk of information-based trading is lower for more active stocks. Surprisingly, information based trading is apparently less common on the recently created Prague Stock Exchange than on more well-established markets.

Abstrakt: Tento článek se zabývá vztahem mezi likviditou a obchodováním založeným na neveřejné informaci v prostředí aukce řízené příkazy na nákup a prodej. K analýze frekvencí příchodu nových informačních signálů a jejich vlivu na strukturu podaných obchodních příkazů je použit model, jehož struktura vychází z článku Easley a kol. (1996). Pomocí tohoto modelu je následně odhadnuta pravděpodobnost obchodování založeného na neveřejné informaci pro soubor akcií obchodovaných na Burze cenných papírů Praha (BCPP). Empirické výsledky souhlasí s předchozími studiemi v tom, že pravděpodobnost obchodování založeného na neveřejné informaci je nižší pro likvidnější akcie. Překvapivým zjištěním je, že obecně pravděpodobnost obchodování založeného na neveřejné informaci je na centrálním trhu BCPP významně nižší než na zavedených kapitálových trzích (New York Stock Exchange).

Keywords: liquidity, information, informed trading, Prague Stock Exchange.

Acknowledgments: I am indebted to Randall Filer and Jan Hanousek for their useful comments and suggestions.

1. Introduction

The connection between liquidity and information based trading has been studied by several authors (see, for example, Hasbrouck (1988, 1991) and Easley, Kiefer, O'Hara and Paperman (1996)) using data from developed capital markets. The main purpose of those investigations was to explain the observed differences in spreads for active and infrequently traded stocks. Easley et al. (who used a sample of stocks listed at the New York Stock Exchange) found that the probability of information-based trading is lower for high volume stocks. Information-based trading also explained, at least partially, the differences in spreads for active and infrequently traded stocks. I ask whether similar conclusions hold in an emerging market such as that of the Czech Republic.

The institutional structure of the Prague Stock Exchange (PSE) is quite different from that of the NYSE. There are no market makers setting bid-ask spreads. Instead, trading is done through the so-called Automated Trading System, which for each stock listed on the market[1], clears the orders made by individual brokers. This clearing is done once a day (at about 11am), and although some additional trading later in the day is possible, this must be done at a price set by the clearing algorithm. The orders submitted by the brokers can be both market (a simple buy/sell order) and limit orders.

Despite these major differences in the institutional structure of these markets (PSE versus NYSE), an econometric model similar to that used by Easley et al. can be utilized to examine the importance of information based trading on the Prague Stock Exchange. Naturally, the purpose of this investigation is not to study the determinants of a bid-ask spread, but rather to inspect the role which informed agents assume in an institutionalized market in an economy where they also have other possibilities to use their information.

These possibilities can be described as follows: In the U.S. (or any other country with a developed capital market), although a particular stock can be traded on several markets, each of those markets should produce a price for the stock which is easily observable by other agents on the market. Possible arbitrage opportunities are then quickly exploited so that differences in prices on different markets disappear or fall below the level of transaction costs connected with arbitrage trading. In the Czech Republic, there exist three points of stock trading: 1) the Prague Stock Exchange, where agents can either trade their stocks on the so-called "central market" (described above

[1] with the exception of a very small group of the most active stocks which also trade continuously through the so-called KOBOS system.

and studied in this paper) or take part in block or direct trading done with larger packages of stocks; 2) the RM-System, a computerized network which also uses an order-clearing mechanism, and due to its wide availability, serves mainly small stockholders; and 3) the Center of Securities, which actually serves as a register of all securities and their holders. Any two agents can trade their stocks directly at the counter of the Center of Securities at any agreed upon price.

Before the beginning of 1995, records of these transactions were not available to other agents. After this date, weekly summaries of volumes and average volume-weighted prices were published. Only recently (from February 1997) have records on all transactions and all prices become available to the public. This means that, for quite a long time, informed agents could utilize their information without revealing it to other traders on the (organized) market simply by exchanging their shares at the Center of Securities rather than on the PSE or RM-System (provided they themselves found a counterpart for this trade). Small capitalization of the PSE and lagged reporting of trading at the Center of Securities make revealing of new information through trade at the PSE more probable than through trade at the Center of Securities, despite the fact that anonymous trading is only possible at the PSE. Given these circumstances, it is reasonable to expect that the role of informed trading on the PSE is generally very low. To test this prior expectation, a model explicitly representing the possibility of information-based trading is presented and the probability of information events and informed trading are estimated for a subsample of PSE stocks.

The rest of the paper is organized as follows: in Section 2, the model is developed and a likelihood function to be maximized is derived. In Section 3, the data are described. Section 4 provides results of the estimation and Section 5 concludes the paper.

2. The Model

2.1 Trading Mechanism of the PSE

The structure of the model is determined by the trading mechanism of the PSE and the type of data that are publicly available. It is important, therefore, to describe the trading procedure in more detail before presenting the model. As already mentioned, trading is done and a price is set by the Automated Trading System (ATS), which clears the buy and sell orders[2] for each stock. The goal of the clearing procedure is to maximize the number of shares traded. In this setting, where shares trade via order-driven auctions, the role of the market versus limit orders is quite different than in a continuous trading framework. In the continuous auction, the limit orders "waiting" in the limit-order book of the market makers provide liquidity to the market and can be matched with the market orders as the latter arrive at the market. However, in the clearing mechanism of the type described above, the market orders are those which, by matching any limit order, increase the volume of trade and liquidity of the stock.

A crucial feature of the price-setting mechanism at the PSE is the upper limit on the percentage price change: for most issues, the price can change by at most 5% during a single trading day (auction). Comparison of demand and supply patterns during the price-setting process can, therefore, have several qualitatively different outcomes. These are summarized in the variable (published by PSE) called "code of the market," which can take one of eight values. These are described in the Appendix. Values of this "code of the market" variable, together with values of the allocation ratio variable (capturing the extent of order rationing) can be used to reconstruct the numbers of shares demanded/supplied in a way that is discussed more deeply in the next section, where the structure of the model is presented.

2.2 The Structural Setup of the Model

The observable pieces of information, namely the degree of demand/supply imbalance measured by the "code of the market" variable and the number of shares demanded/supplied at the new market price, can be utilized in the following model to provide insights into trade and information flows.

Traders arriving at the market can either be uninformed or have information about the value of the stock. On each day, prior to the beginning of the period in which

[2] The order may have a limit price specified, i.e., a maximum price for buy and minimum price for sell order. If the limit price is not specified, it is a simple market order.

orders can be submitted to the market, nature determines whether an information event relevant to the value of the particular asset will occur. These information events are assumed to be independently distributed and to occur with probability α. There are good signals with probability 1 - δ, and bad signals with probability δ. Informed investors know these signals while uninformed ones do not.

As Easley et al. (1996), I assume that the arrival of uninformed buyers and sellers at the market (i.e., their decision to trade and the actual submission of orders — to be considered in the upcoming auction — to the market) is determined by an independent Poisson process with the arrival rate ε. The arrival of news (signals) to the traders and their subsequent arrival at the market are assumed to follow a Poisson process with the arrival rate μ. Observing a good (bad) signal leads to the submission of a buy (sell) order. This "informed trading" process is also assumed to be independent of the arrival processes of uninformed traders.

The described Poisson processes result from each agent deciding (not having information or observing a signal) whether to trade at the market or not. They do not capture the decision about the type of order (market versus limit) and, in the case of a limit order, the decision about limit price. In fact, I assume that this decision is independent of the decision on whether or not to trade, and I thus model it with the *ad hoc* construction described below.

If P_{t-1} is the previous-day price of the stock and P^*_t is the new price, define the price change $p^* = (P^*_t - P_{t-1})/P_{t-1}$ and measure the price of the stock on a given day by the relative price change p^*. Because of limits on price change due to the trading mechanism of the PSE, a trade could never occur at a price outside the admissible (5%) interval. We can, therefore, think of market orders as a special type of limit order with the limit price set 5% below (above) the previous-day price in the case of sell (buy) orders. The limit prices of particular orders submitted are discretely distributed in the admissible interval. This is approximated here by the (continuous) logistic function, which specifies number of shares demanded at price p as

$$B(p) = A^D \, \frac{e^{-\beta(p+s)}}{1+e^{-\beta(p+s)}}\,, \tag{1}$$

where β and s are parameters, and A^D is the total number of shares demanded — irrespective of possible limit prices of corresponding buy orders — which was assumed above to follow a Poisson distribution.

Similarly, the number of shares supplied at price p is modeled as

$$S(p) = A^S \frac{e^{\beta(p-s)}}{1+e^{\beta(p-s)}}, \qquad (2)$$

where A^S is the total number of shares supplied — irrespective of the corresponding limit prices — which also follows a Poisson distribution.

Depending on the relative position of B(p) and S(p) (driven by the realization of A^D and A^S), there are three possible outcomes:

a) Demand and supply patterns cross for p^*, which lies inside the admissible region. This corresponds to "codes of the market" 1, 2, and 3 in the real data.

b) The new price is at the boundary and the allocation ratio is greater or equal to 20%. This corresponds to "codes of the market" 4 and 5 in the data.

c) The allocation ratio for p^* at the boundary of the admissible region is lower than 20%, or there are no orders submitted on either the buy or sell side of the market. This means that no trade is made and corresponds to "codes of the market" 6, 7, and 8 in the real data.

Codes 1, 2, and 3 correspond to the situation in which the market price p^*, proposed by the ATS, falls inside the admissible interval, and all orders valid at that price are executed. In that case, the quantity (number of shares) demanded, $B(p^*)$, is equal to the quantity supplied, $S(p^*)$. If the new price lies at the boundary (codes 4 and 5), some orders may be rationed. The extent of rationing is given by the allocation ratio discussed in the Appendix and this ratio (as published by PSE) can therefore be used to reconstruct the quantities $B(p^*)$ and $S(p^*)$ from the number of actually exchanged shares.

Given the parameters β and s, the number of shares demanded/supplied at the new price p can be used to compute the (unobservable) total amounts demanded (A^D) and supplied (A^S). While it is possible to reconstruct these numbers for cases a) and b) corresponding to the "codes of the market" 1, ..., 5, it is not possible to infer them from the information published in the case of no trade ("codes of the market" 6, 7, and 8).

There are several ways to deal with this problem. First, one can simply condition on nonzero trade volume and restrict oneself to the "codes of the market" 1, ..., 5. Second, given the structural setup of the model, it is *in principle* possible to compute the probability of no trade (as a function of structural parameters) explicitly from the assumptions on demand/supply patterns and to use the observations with "codes of the market" 6, 7, and 8 to obtain more efficient estimates of the parameters of the model. This would, however, involve computation of quantiles of the distribution of the ratio of two Poisson random variables, which would be extremely cumbersome.

The approach used here lies somewhere between the first case of abandoning the additional information contained in observations of no trade and the second case of explicit (but very complicated) modeling.

The approach utilizes the simple implication of the model for demand/supply patterns as described in (1) and (2); namely, that in the case of an information event, the probability of "global mismatch" between demand and supply (resulting in the case of no trade) should be higher than in the case when no new information occurs. Thus, if γ_E is the probability of nonzero trade in the case of a new signal, and γ_N denotes the probability of nonzero trade when there is no new information, the inequality $\gamma_N \geq \gamma_E$ should hold. This restriction is incorporated into the model in an attempt to utilize the observations of no trade, particularly for better estimation of the probability of an information event, α.

Another justification for this *ad hoc* treatment of the fact that A^D and A^S are unobservable in the case of no trade is to allow more flexibility in the Poisson-type model, which could be too restrictive with respect to the probability of no trade. Particularly, the probability of no trade as implied by the model specification could be too low compared to empirical evidence. Therefore, the much easier estimation and higher flexibility of the simplified model would seem to outweigh the possible efficiency loss due to this simplification. The structure of the proposed model as discussed in the preceding paragraphs is illustrated in Figure 1.

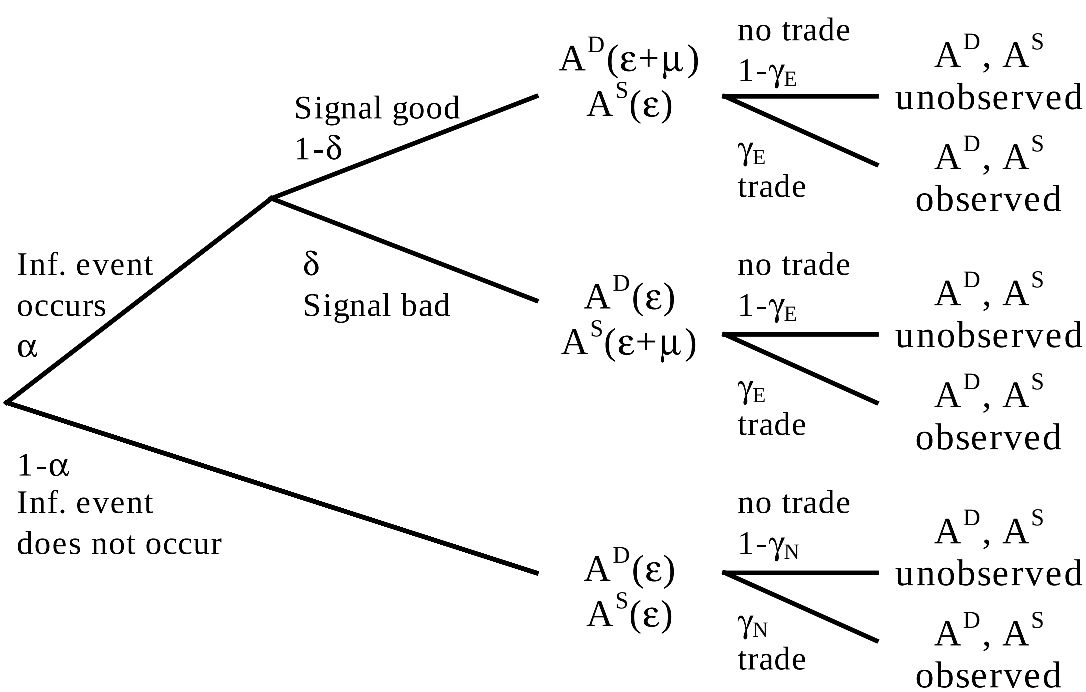

Figure 1: **Tree diagram of the trading process.** α is the probability of an information event, δ is the probability of a bad signal, ε is the arrival rate of uninformed traders, μ is the arrival rate of informed traders, γ_E is the probability of nonzero trade in the case of an information event, and $\gamma_N \geq \gamma_E$ is the probability of nonzero trade when no new information occurs. $A^D(\omega)$ ($A^S(\omega)$) indicates that the total number of shares demanded (supplied) follows a Poisson process with the arrival rate ω. These values are unobserved if there is no trade and can be reconstructed (using (1) and (2)) from the data on shares demanded/supplied at the new price, otherwise.

The parameters to be estimated for a given stock are α, δ, ε, μ, γ_E, γ_N, β and s. The primary goal of the paper is, however, to estimate the extent of information based trading. Based on the branch of the tree prevailing on a given day, buy and sell orders follow different arrival processes. The "average arrival rate" of informed traders is $\alpha((1-\delta)\mu + \delta\mu) + (1-\alpha)*0 = \alpha\mu$. The "average arrival rate" of all traders is $\alpha((1-\delta)(\mu+2\varepsilon) + \delta(\mu+2\varepsilon)) + (1-\alpha)*2\varepsilon = \alpha\mu+2\varepsilon$. The overall probability of informed trading is given by the ratio of these two expressions, so that

$$PI = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon}. \tag{3}$$

2.3 The Likelihood Function

In the proposed model buy and sell orders follow one of three Poisson processes on each day. Whether or not new information occurs is not directly observable. It is, however, reflected in the data so that more buy orders are expected on good-signal days, and more sell orders are expected on bad-signal days. On no-event days, there are no informed traders arriving at the market and fewer trades can be expected. The probabilities of these cases are determined by the probability of new information occurring and the type of information.

To construct the likelihood function for the whole model, the likelihood of order arrivals on a day of known type is derived first. Consider a good-signal day. The buy orders arrive at rate $\mu + \varepsilon$ as both uninformed and informed traders submit these orders. The sell orders arrive at rate ε as only uninformed traders sell. The distributions of the total number of shares supplied/demanded are independent Poisson distributions. Then, the likelihood of observing the total of A^D buy orders and A^S sell orders submitted on a good-event day, conditional on nonzero trade, is

$$e^{-(\mu+\varepsilon)} \frac{(\mu+\varepsilon)^{A^D}}{A^D\,!} e^{-\varepsilon} \frac{\varepsilon^{A^S}}{A^S\,!}. \tag{4}$$

Similarly, on a bad-event day, this conditional likelihood is

$$e^{-\varepsilon}\,\frac{\varepsilon^{A^D}}{A^D\,!}\,e^{-(\mu+\varepsilon)}\,\frac{(\mu+\varepsilon)^{A^S}}{A^S\,!}\,, \tag{5}$$

and for a no-event day, it is

$$e^{-\varepsilon}\,\frac{\varepsilon^{A^D}}{A^D\,!}\,e^{-\varepsilon}\,\frac{\varepsilon^{A^S}}{A^S\,!}\,. \tag{6}$$

Let D_{NT} denote the "no trade" 0-1 indicator variable, which is 1 if there was no trade with the stock on a given day (code of the market equal to 6, 7 or 8), and 0, otherwise. The probabilities of a good-event day, bad-event day, and a no-event day are $\alpha(1-\delta)$, $\alpha\delta$, and $1-\alpha$, respectively. The overall likelihood function for a given day, therefore, is

$$\begin{aligned} L((A^D, A^S)|\theta) &= D_{NT}\left[\alpha(1-\gamma_E)+(1-\alpha)(1-\gamma_N)\right] \\ &+(1-D_{NT})\left[\begin{aligned}&\alpha(1-\delta)e^{-(\mu+\varepsilon)}\,\frac{(\mu+\varepsilon)^{A^D}}{A^D\,!}\,e^{-\varepsilon}\,\frac{\varepsilon^{A^S}}{A^S\,!} \\ &+\alpha\delta e^{-\varepsilon}\,\frac{\varepsilon^{A^S}}{A^S\,!}\,e^{-(\mu+\varepsilon)}\,\frac{(\mu+\varepsilon)^{A^S}}{A^S\,!}+(1-\alpha)e^{-\varepsilon}\,\frac{\varepsilon^{A^D}}{A^D\,!}\,e^{-\varepsilon}\,\frac{\varepsilon^{A^S}}{A^S\,!}\end{aligned}\right] \end{aligned} \tag{7}$$

The values of A^D and A^S are not directly observable and are modeled here by (1) and (2), from which we can write

$$A^D = \left\lfloor B\left[1+e^{\beta(p+s)}\right]\right\rfloor, \tag{8}$$

$$A^S = \left\lfloor S\left[1+e^{-\beta(p-s)}\right]\right\rfloor, \tag{9}$$

where B and S are the number of shares demanded/supplied at the new market price, and $\lfloor x \rfloor$ denotes the nearest integer to x; this rounding is employed as A^D and A^S are assumed to follow (discrete) Poisson distribution. Substituting from (8) and (9) into (7), we get the likelihood $L((B,S)|\theta)$ in terms of observable variables. As days are independent, the likelihood of observing the data $(B_i, S_i)^I_{i=1}$ over I days is the product of daily likelihoods,

$$L = \prod_{i=1}^{I} L(\theta|(B_i, S_i))\,. \tag{10}$$

This function is then maximized to estimate the parameter vector θ.

3. The Data

The model derived in the previous section is estimated for a subsample of stocks traded on the PSE. All data available from the introduction of a given stock on the market[3] until November 30, 1996 are used. For each stock in the sample, we need to estimate the parameters of the trade process. If the stock is traded too infrequently, there may not be enough data for this estimation. Also, the price level of a given stock can influence the trading process. These issues are discussed below where I describe the sample selection criteria.

The sample of all joint stock companies traded on the PSE was first sorted on the average probability of nontrading (probability of no trade in a given day). The sample is then divided into deciles, where the first decile contains the most frequently traded stocks. Although there were more than 1,750 firms listed on the PSE in 1996, most of them did not trade very frequently (see Němeček (1996) for details). Trading frequencies and volumes decrease rapidly across deciles. To be able to judge the role of trading activity, and at the same time to be able to estimate the model[4], I use stocks from the first through the fifth nontrading deciles.

To eliminate the possible effects of the stock price levels, I construct a matched sample of stocks with the "same" prices but different trading frequencies. This matching is very similar to the one used by Easley et al.: the average price of the stock over the whole period was used to sort the stocks within the nontrading deciles. The adjacent pairs of stocks from different deciles are then matched into groups. This yields 175 groups out of which 25 groups are randomly selected.

Table A.1 in the Appendix lists the group statistics for the selected stocks, their average probabilities of nontrading, book values of the firms (number of shares outstanding times their face value), average prices, and the number of observations (number of days on the market). The individual stock data (not shown, available from author) show clearly that the extent of nontrading, as well as the average price of the stock, is closely related to the size of the firm. Larger firms have a much higher liquidity and a higher average price.

4. Estimation

In section 2, the likelihood function for the structural model was derived. This likelihood function can be maximized, conditional on trade data for a given stock, to obtain the estimates of the trade process and information flow for that stock. The

[3] June 22, 1993 and March 1, 1995 for most of the stocks offered in the first and second wave of the voucher privatization, respectively. The vast majority of stocks entered the market on one of those two dates.
[4] The estimation process often does not converge for infrequently traded stocks.

probability parameters α, δ, γ_E, and γ_N were restricted to (0,1) by a logit transformation of the unrestricted parameters. To ensure that $\gamma_N \geq \gamma_E$, as predicted by the model, γ_N was expressed as $\gamma_E + \gamma$, and this summation was also restricted to (0, 1). The arrival-rate parameters ε and μ were restricted to $(0, \infty)$ by a logarithmic transformation. The likelihood function expressed in terms of these unrestricted parameters was then maximized using the ML procedure of the TSP package. Standard errors for the economic parameter estimates were calculated from the asymptotic distribution of the unrestricted parameters using the delta method.

Because of their length, detailed results of the estimation are not presented here. However, to illustrate the individual-firm results and complete the picture as presented by group statistics reported in Table 1 and discussed below, the estimated parameters and corresponding standard errors for each stock in the middle-nontrading group C of the sample are provided in Table A.2 of the Appendix. The values of t-statistics reported in Table A.2 show that the parameters of the model are, in most cases, estimated quite precisely; this is true especially for the arrival rates of uninformed (ε) and informed (μ) traders.

Note that the estimates of parameters β and s of the demand and supply functions are not reported. The reason is that these parameters were never significantly different from zero. There are several possible explanations for this observation. First, if the limit prices of all orders lay at the boundaries of the admissible new-price interval so that these orders were valid throughout the whole interval (and they were, in fact, equivalent to the market orders with no limit price specified), the demand and supply patterns would be flat, justifying an estimate of β equal to zero. However, this does not seem to be an accurate account. A second explanation is that the model chosen for the shape of demand and supply patterns inside the admissible region (the logistic function) may not be appropriate. I have, therefore, tried a simpler model using a linear structure for these patterns. The estimates of this linearized model in no way differ from those reported. The third, and most probable, explanation is that the model as designed requires simply too much from the data, and therefore, it is not possible to reliably estimate the shape of the demand and supply patterns from the aggregate number of buy and sell orders on any given day. Having some special information about the type of day (whether new information occurred and what type of information it was) could help. If there are signals that are not purely firm-specific but influence several firms simultaneously, then the application of panel data techniques could help to resolve this problem. I will return to this point in the conclusion.

<u>Table 1:</u> Summary statistics on estimated parameters by group

Parameter:	PIT (%)	ε	μ	α (%)	δ (%)
A: First decile					
Mean	0.19	787.266	2.590	13.53	20.75
Median	0.00	319.892	0.000	0.01	22.70
Std. dev.	0.32	1091.865	3.861	17.35	13.95
Stat. sign. at 5%	57.14%	100.00%	56.25%	44.44%	12.00%
B: Second decile					
Mean	0.74	122.53	3.76	30.80	12.06
Median	0.65	102.621	4.562	30.80	2.84
Std. dev.	0.77	74.71	2.87	15.35	14.30
Stat. sign. at 5%	68.18%	100.00%	72.73%	65.22%	8.33%
C: Third decile					
Mean	1.39	114.989	5.316	28.95	11.75
Median	0.88	76.474	4.219	30.49	5.53
Std. dev.	1.72	124.745	3.916	11.90	13.58
Stat. sign. at 5%	56.52%	100.00%	70.83%	62.50%	22.73%
D: Fourth decile					
Mean	1.93	72.68	6.46	26.17	9.41
Median	1.50	66.621	6.835	23.02	0.00
Std. dev.	2.56	51.23	3.91	15.65	15.94
Stat. sign. at 5%	36.00%	100.00%	83.33%	40.00%	16.67%
E: Fifth decile					
Mean	2.01	64.377	7.069	19.29	18.67
Median	0.81	46.688	6.401	14.72	7.88
Std. dev.	2.59	56.559	5.231	12.79	23.72
Stat. sign. at 5%	58.33%	95.24%	93.33%	60.00%	25.00%

To see the differences in estimated parameters among the groups, group statistics were computed (<u>Table 1</u>). Besides the mean, median, and standard deviation, this table shows a percentage of firms in a given group, for which an estimate of a given parameter is statistically significant at the 5% significance level. Though the ranking of the estimates of probability of informed trading and arrival rates across deciles is consistent with prior expectations, the variability of the estimates suggests that there are no significant differences in means of the parameters across groups[5].

Like Easley et al., I, therefore, compute nonparametric statistics to compare the distributions of estimated variables across groups. Specifically, the Kruskal-Wallis test and the Wilcoxon rank sum test are used. The Kruskal-Wallis test serves to check whether the five population distribution functions are identical against the alternative that at least one of them is different. The Wilcoxon test is used for a pairwise comparison of these distribution functions and tests whether the values for one sample

[5] This fully agrees with the findings of Easley et al., who also find a similar ranking of not significantly different group means.

tend to be higher or lower than for the second sample. The values of the test statistics are given in Table 2.

Consider first the parameters capturing the arrival rates of uninformed and informed traders. Table 1 shows large differences in both ε and μ across groups. Naturally, the arrival rate of uninformed traders falls dramatically when going from the most active stocks to less liquid ones. The Kruskal-Wallis test rejects the hypothesis that these arrival rates are equal across groups; the Wilcoxon test supports this observation by showing that the distribution of ε for the most active stocks (A) differs significantly from that of other groups — B, C, D and E. Only the differences between the adjacent groups B and C, C and D, and D and E are not significant.

Table 2: Nonparametric tests

The Kruskal-Wallis test is used to determine whether the five populations (groups of stocks) from which the parameter values are drawn are identical against the alternative that at least one of them differs. The Wilcoxon rank sum test compares two populations and indicates whether one of them tends to yield higher values. Deciles are denoted by A (most frequently traded stocks) through E (fifth nontrading decile).

Parameter:	PIT	ε	μ	α	δ
Kruskal - Wallis Tests					
	31.692	52.474	22.195	24.244	13.074
	Asymptotically distributed as $\chi^2(4)$.				
	Critical value for $\alpha = 0.05$ is 9.49				
Wilcoxon rank sum test					
A - B	3.328	-4.472	2.066	3.502	-2.260
A - C	3.754	-4.686	2.862	3.386	-2.202
A - D	4.763	-5.384	3.619	2.590	-3.250
A - E	4.084	-5.501	3.328	1.775	-1.387
B - C	1.407	-1.426	1.348	-0.223	0.126
B - D	2.474	-2.571	2.454	-2.086	-0.922
B - E	1.659	-3.716	2.280	-2.823	1.310
C - D	1.116	-1.426	1.019	-1.659	-1.019
C - E	0.572	-2.571	1.019	-2.998	0.999
D - E	-0.301	-0.728	0.049	-2.028	2.163
	Asymptotically distributed as N(0,1).				
	Critical value for $\alpha = 0.05$ is +-1.645				

The distributions of arrival rates of informed traders also exhibit the expected patterns. According to Table 1, the arrival rate of informed traders increases when going from more active to less active stocks, not only in relative terms (compared to the arrival rate of uninformed traders), but also in absolute values.[6] Again, both the Kruskal-Wallis test and Wilcoxon test confirm this ranking and show significant differences between most of the deciles. The hypothesis of identical distribution is not

[6] This is different from the observation of Easley et al., who find that (in absolute terms) the arrival rate of informed traders is higher for more active stocks.

rejected for pairs B and C, C and D, C and E, and D and E, i.e., primarily for the groups of less traded stocks.

Unlike in Easley et al., the estimates of the information event parameters α and δ are not linearly related to the stocks' liquidity. While the probability of an information event occurring on a given day has a mean of 13.53% for the first decile, it increases to 30.80% for the second decile and falls back to 19.29% for the fifth decile. Thus, only the low mean probability of an information event occurring for the first decile does not conform to the pattern of decreasing frequency of information events for less active stocks found in Easley et al. Also, the magnitudes of these frequencies are significantly lower than those found for NYSE stocks (about 50% for the first and 35% for the eighth decile).

Moreover, the pattern of good/bad signals is reversed — the probability of a bad signal falls from 20.75% for the most active stocks to 9.41% for the fourth group and increases to 18.67% for the fifth decile of PSE stocks; it increases from about 35% for the first decile to 50% for the eighth decile of NYSE stocks.[7] As for the statistical significance of these observations, the Kruskal-Wallis test leads to rejection of the hypothesis of an identical distribution of parameter α across groups; this test also finds (at 5% level) significant differences in the distributions of types of signals (parameter δ) across the five groups. The Wilcoxon rank sum test finds statistically significant differences in the distributions of parameter α for almost all pairs of groups — the only exception is the pair B and C, i.e., the second and third nontrading deciles. The test also supports the hypothesis of a higher probability of a bad signal (δ) for the most active stocks, compared to other groups.

Finally, consider the parameter which is of highest interest — the overall risk of informed trading. The group means reported in Table 1 show an increase in this parameter when going from more active to less active stocks. Both the Kruskal-Wallis and Wilcoxon tests indicate significant differences in the distributions of the probability of informed trading across groups. Only when the bottom three deciles are compared does the Wilcoxon test not reject the hypothesis of identical distributions. This again

[7] There is no clear explanation of the relatively low probabilities of the bad signals on the PSE. One possibility is, however, that because of the concentration process, the bad signals are reflected more in a decrease in the trading activity than in an increase in the supply of shares (investors who want to acquire a higher portion of the firm will be more reluctant to sell its shares when observing a bad signal). This would mean that arrival rates of informed traders in the case of an information event are different for good and bad signals. The model as estimated did not allow for such a possibility.

agrees with the observation of Easley et al. that probability of informed trading increases with decreasing liquidity.

5. Conclusions

Recent discussions about the connection between liquidity and the informational role of stock markets yield important insights into the market makers' decision on the bid-ask spreads in the framework of continuous-time dealer markets. A clear conclusion of empirical studies is that low liquidity enhances the risk of information-based trading and that wider spreads for infrequently traded stocks do — at least partially — result from market makers insuring themselves against losses from trading with an informed agent.

This paper tries to access the same problem in the setting of order-driven auction markets. For a sample of Prague Stock Exchange stocks, the behavior of active and infrequently traded stocks was investigated. A model, similar in spirit to that of Easley et al., but taking into account the limits of data availability and other specifics of the PSE, was proposed and the probability of information-based trading was estimated for each stock in the sample. While it was not possible to estimate the underlying shapes of the demand and supply patterns, the arrival rates of uninformed and informed traders and probabilities of information events were estimated with reasonable precision. The analysis confirms the *a priori* expectation that the probability of information-based trading is lower for actively traded stocks than it is for inactive stocks.

The relative patterns and observations on the role of liquidity in determining the extent of information-based trading, reported above for the case of the order driven (auction) stock trading at the PSE, are in full accord with the findings of Easley et al. for the price driven (continuous auction) trading at NYSE. The extent of informed trading, however, differs substantially. While Easley et al. report average probabilities of informed trading of about 20%[8], the average risks of information-based trading found here are just 0.19%, 1.39%, and 2.01%, respectively, for the corresponding three groups (deciles) of PSE stocks. There are several factors that could explain this difference.

First, as discussed in the introduction, the situation in the Czech Republic in the analyzed period was special in the sense that informed traders had many opportunities to utilize their information outside the price-making market. This could also explain why the arrival rates of informed traders for active stocks are lower in magnitude than the arrival rates of informed traders for inactive stocks — it is easier for an informed trader to find another party for a transaction that involves actively traded stocks than it is to find one for trading an illiquid stock. However, when an informed agent trades directly with another party (outside the market), other agents may more easily identify him as being informed. In the case of a developed market, where all agents are price-takers on the market, an informed trader is more likely to trade in the market than outside because of the risk of revealing his type. The PSE, however, is a market with quite small capitalization, one in which orders submitted by one agent may substantially influence the auction price of the particular stock. At the same time, there was (during the period studied in this paper) a substantial time lag in reporting of the transactions made at the Center of Securities and the reporting requirements were very weak. Further, the transaction costs of the transfer at the counter of the Center of Securities were negligible compared to the costs of trading on the market — this difference in trading costs was, of course, largest for the most active stocks (which have, at the same time, the highest prices). These facts are likely to have led the informed agent to trade primarily outside the market.

Another argument explaining the low average probabilities of informed trading comes from the work of Pagano and Röell (1996), who show, in a stylized environment, that greater transparency of the trading mechanism reduces trading costs and the ability of informed traders to profit from their information. In this respect, the order-driven auction of the PSE ranks above the continuous auction of the NYSE. This also supports the lower risk of informed trading on the PSE.

To conclude, the proposed model documents that the risk of informed trading decreases with increasing liquidity of a stock. The extent of information-based trading at the PSE is very low in magnitude, which can be explained by the other opportunities informed traders have to utilize their private information and by the institutional structure of the market. The shapes of demand and supply patterns which could not be identified in this model may, hopefully, be revealed when panel data techniques are applied to enhance the model and to utilize information carried by signals that are not purely firm-specific.

[8] 16.4%, 20.8%, and 22.0% for the three groups they use

REFERENCES

Diamond, D. W., Verrecchia, R. E. (1991). “Disclosure, Liquidity, and the Cost of Capital,” *Journal of Finance* 46(4): 1325 - 1359.

Easley, D., Kiefer, N. M., O'Hara, M. (1996). “Cream-Skimming or Profit-Sharing? The Curious Role of Purchased Order Flow,” *Journal of Finance* 51(3): 811 - 833.

Easley, D., Kiefer, N. M., O’Hara, M., and J. B. Paperman (1996). “Liquidity, Information, and Infrequently Traded Stocks,” *Journal of Finance* 51: 1405 - 1436.

Hasbrouck, J. (1988). “Trades, Quotes, Inventory, and Information,” *Journal of Financial Economics* 22: 229 - 252.

Hasbrouck, J. (1991). “Measuring the Information Content of Stock Trades,” *Journal of Finance* 46: 179 - 207.

Němeček, L. (1996). “Expected Returns When Shares Are Subject to Infrequent Trading: Evidence from the Czech and Slovak Stock Markets,” *unpublished*.

Pagano, M. and A. Röell (1996). “Transparency and Liquidity: A Comparison of Auction and Dealer Markets with Informed Trading,” *Journal of Finance* 51: 579 - 611.

APPENDIX

Description of the variable “code of the market” published by the PSE:

Code = 1 ... *perfect balance*
Number of securities supplied and number of securities demanded at the new price are equal. The new price is within the allowed spread margin.

Code = 2 ... *local excess on supply side*
Number of securities supplied at the new price is higher than the number of securities demanded at that price. The new price is within the allowed spread margin. All buy orders with limit price higher or equal to the new price will be satisfied. Sell orders with limit price lower or equal to the new price will be rationed[9].

Code = 3 ... *local excess on demand side*
Number of securities demanded for purchase at the new price is higher than the number of securities supplied at that price. The new price is within the allowed spread margin. All sell orders with limit price lower or equal to the new price are satisfied. Buy orders with limit price higher or equal to the new price will be rationed[3].

Code = 4 ... *global excess on supply side*
Number of securities supplied at the new price is higher than the number of securities demanded at that price. The new price equals the lower limit of the allowed price spread margin. All buy orders with limit price higher or equal to the new price will be satisfied. Sell orders with limit price lower or equal to the new price will be rationed according to the allocation ratio, which is the ratio of shares actually sold to the number of shares supplied.

Code = 5 ... *global excess on demand side*
Number of securities demanded at the new price is higher than the number of securities supplied at that price. The new price equals the upper limit of the allowed price spread margin. All sell orders with limit price lower or equal to the new price will be satisfied. Buy orders with limit price higher or equal to the new price will be rationed according to the allocation ratio, which is the ratio of shares actually sold to the number of shares demanded.

Code = 6 ... *total excess on supply side*
Number of securities supplied at the new price is higher than the number of securities demanded at that price. The new price equals the lower limit of the allowed price spread. The allocation ratio is lower than 20 %, and therefore no trade will take place.

Code = 7 ... *total excess on demand side*
Number of securities demanded at the new price is higher than the number of securities supplied at that price. The new price equals the upper limit of the allowed price spread. The allocation ratio is lower than 20 %, and therefore no trade will take place.

Code = 8 ... *not quoted*
No sell or buy orders were placed, or such orders were placed, but the limit prices of sell orders do not overlap with the limit prices on the demand side. The previous price remains valid. No trade takes place.

[9] This could happen due to the discrete nature of the distribution of orders actually submitted. In the model, the stepwise non-increasing (non-decreasing) functions describing cumulative numbers of shares demanded (supplied) at a given price were modeled by the smooth logistic functions so that rationing occurs only at the boundaries of the admissible price interval.

<u>Table A.1:</u> Data on groups of PSE stocks included in the sample

This table presents group statistics of the data on stocks included for estimation. Book value is measured in thousands of CZK.

	Prob. of nontrading	Book value	Average price	# of obs.
A: First decile				
Mean	7.31%	2320094	1045.64	596
Median	6.65%	1033552	546.46	654
Std. dev.	2.56%	2852227	1201.57	96
B: Second decile				
Mean	22.37%	528341	479.04	541
Median	22.57%	463704	220.82	506
Std. dev.	5.73%	305027	603.04	112
C: Third decile				
Mean	41.03%	343051	364.92	550
Median	39.92%	282858	192.94	651
Std. dev.	4.60%	281362	400.05	112
D: Fourth decile				
Mean	57.32%	183022	290.50	543
Median	57.26%	125807	192.47	535
Std. dev.	3.05%	157732	262.28	111
E: Fifth decile				
Mean	67.06%	116531	205.36	548
Median	66.67%	101963	136.61	654
Std. dev.	2.28%	91529	166.04	114

<u>Table A.2:</u> Parameter estimates for the middle-nontrading group C

This table presents the estimates of the parameters of the model for each stock included in the middle-nontrading group C of the sample. BIC stands for Bohemian Identification Code, an official tag used by the PSE authorities for a given firm. PIT gives the probability of informed trading as derived from the other parameters: ε, arrival rate of uninformed traders, μ, arrival rate of informed traders, and α, probability of an information event. δ gives the probability of a bad signal; γ_E (γ_N) the probability that there will be trade, conditional on an information event occurring (not occurring). Values of t-statistics are given in parentheses below the parameter estimates. A maximum likelihood estimation was performed using the ML procedure of TSP package. "-" marks the cases where the estimation algorithm failed to provide estimates of standard errors.

BIC	PIT (%)	ε	μ	α (%)	δ (%)	γ_E (%)	γ_N (%)
Group C: Third nontrading decile stocks							
BAATSLHK	0.24	350.160	5.023	33.84	35.89	51.19	71.39
	(5.22)	(4269.35)	(37.63)	(5.59)	(4.63)	(11.91)	(28.34)
BAABUZUL	0.64	71.019	3.797	24.24	5.53	62.74	71.63
	(3.43)	(936.37)	(27.36)	(3.68)	(0.13)	(21.08)	(31.60)
BAAAVIAB	2.52	82.661	11.472	37.23	16.09	53.19	64.80
	(7.98)	(737.49)	(64.88)	(7.78)	(2.21)	(8.80)	(14.38)
BAAPEGA	1.42	71.274	6.520	31.53	0.00	42.80	55.40
	(0.92)	(181.63)	(2.08)	(0.67)	-	(4.93)	(8.18)
BAAFASAD	0.88	58.181	3.927	26.36	0.00	58.69	69.66
	(4.96)	(791.83)	(27.68)	(5.44)	(0.00)	(20.50)	(33.13)
BAALIGRA	1.62	148.961	12.378	39.65	0.00	55.52	66.51
	(3.21)	(730.22)	(28.91)	(3.26)	(0.00)	(10.79)	(17.16)
BAAODKOL	0.43	50.484	2.529	17.06	12.78	56.55	68.24
	(0.06)	(481.17)	(19.84)	(0.06)	(1.06)	(0.06)	(0.17)
BAAUNCUK	0.00	148.413	0.000	50.00	0.00	50.00	63.45
	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
BAATRIOD	0.18	211.480	2.860	26.30	0.04	52.03	64.09
	(6.34)	(2937.73)	(27.64)	(6.33)	(0.15)	(7.09)	(17.15)

BIC	PIT (%)	ε	μ	α (%)	δ (%)	γ_E (%)	γ_N (%)
BAAJIHLE	0.24	76.474	1.262	29.71	0.00	52.08	69.73
	(1.83)	(752.42)	(9.43)	(1.90)	-	(11.63)	(24.65)
BAAHOTOT	0.45	98.609	2.932	30.49	0.00	30.49	68.01
	(3.99)	(930.26)	(14.98)	(4.38)	(0.00)	(4.60)	(22.28)
BAARUBEN	0.32	117.975	3.934	19.37	12.03	55.77	71.96
	(3.12)	(1995.92)	(43.12)	(3.13)	(0.88)	(16.49)	(33.57)
BAAGENOS	2.45	44.392	7.022	31.78	0.00	31.78	66.65
	(0.00)	(433.69)	(45.56)	(0.00)	(0.00)	(0.00)	(0.00)
BAAOBTIS	3.29	42.255	11.106	25.91	0.00	56.86	68.46
	(3.84)	(386.70)	(46.93)	(3.80)	(0.00)	(14.27)	(23.51)
BAAMSLZ	0.00	614.493	2.394	0.00	0.01	57.65	58.19
	-	(603.03)	(1.00)	-	(1.00)	(2.62)	(25.81)
BAACADCB	1.40	93.079	7.603	34.68	21.39	34.68	64.60
	(6.13)	(202.13)	-	(6.05)	(2.48)	(6.18)	(10.15)
BAASKRBR	4.11	49.262	11.474	36.81	29.68	36.81	56.30
	(2.18)	(181.11)	(23.73)	(2.16)	(0.59)	(5.40)	(11.95)
BAACALOF	1.39	60.718	5.504	31.05	36.74	31.05	79.68
	(0.29)	(56.20)	(0.36)	(0.21)	(0.07)	(0.22)	(1.94)
BAAJALTA	0.02	54.598	0.050	50.00	26.89	50.00	63.45
	(0.00)	(4.79)	(0.00)	(10.16)	(0.00)	(13.80)	(23.95)
BAACKDKH	1.89	99.874	10.868	35.38	32.59	56.64	59.83
	(12.53)	(853.54)	(52.94)	(12.29)	(3.77)	(19.32)	(39.50)
BAAPRAGL	2.24	51.482	6.350	37.17	11.89	44.65	55.77
	(0.05)	(19.45)	(0.04)	(0.11)	-	(1.00)	(1.57)
BAAPRPLY	1.30	42.790	4.219	26.66	0.00	26.66	91.19
	(5.91)	(563.01)	(26.22)	(6.57)	(0.00)	(5.00)	(142.25)
BAAKOUZE	7.67	16.067	8.755	30.49	23.44	30.49	70.52
	(0.00)	(187.39)	(38.12)	(0.00)	(2.29)	(0.00)	(0.00)
BAASROUT	0.11	77.862	0.912	18.05	0.00	51.24	64.36
	(0.02)	(17.94)	(0.04)	(0.01)	(0.00)	(5.02)	(8.63)
BAAZZNPR	0.00	142.171	0.000	0.00	28.76	0.00	39.41
	-	(216.00)	(1.00)	(1.00)	(1.40)	(1.00)	(23.37)