JEL Classification: G14, G15, P34
Keywords: dealers' market, emerging markets, informed trading, trading systems

Detecting Information-Driven Trading in a Dealers Market*

Jan HANOUSEK – CERGE-EI, Charles University and the Academy of Sciences, Prague; Anglo-American University, Prague; The William Davidson Institute, Michigan; and CEPR, London (jan.hanousek@cerge-ei.cz) – *corresponding author*

František KOPŘIVA – CERGE-EI, Charles University and the Academy of Sciences, Prague; Czech National Bank (frantisek.kopriva@cerge-ei.cz)

Abstract

We focus on the extent of information-driven trading originating from order flows to capture the behavior of the market makers on an emerging market. We modified the classical Easley et al. (1996) model for the probability of informed trading using a jackknife approach in which trades of one particular market maker at a time are left out from the sum of all buys and sells. Using the estimates from the jackknife approach, for each market maker we test whether the order flows associated with the particular market maker behaved significantly differently from the others. Data from the Prague Stock Exchange SPAD trading platform are used to demonstrate our methodology. Finding significant differences in the probability of informed trading computed from order flows, we conclude that order flows could reveal the extent of information-driven trading and could potentially be used by regulatory authorities to identify suspicious behavior by market participants.

1. Introduction

A significant number of studies deal with the issue of insider or informed trading on developed and emerging markets. Starting with the seminal work of Kyle (1985), various models were developed for insider or informed trading and many empirical studies attempted to estimate the severity of this problem. Insider trading can be described as a situation where the investor is trading based on private information that is available only to a restricted number of people. Although insider trading is illegal in many countries, the boundary between insider trading and informed trading is not as obvious as it may look.[1]

To measure the probability of information-driven trading (PIN) Easley et al. (1996) developed a model commonly used in the literature to estimate PIN that is based on the imbalance of buy and sell order flows. Note that PIN is not exclusively an insider trading measure, as it also captures informed trading by investors who are particularly skillful in analyzing public news. It has been shown, for example, by Vega (2006) that the estimated PIN was actually higher after company reports become publicly available. There are two main sources of information: information from firm

* The authors would like to thank Jan Bena, Randall Filer, Evžen Kočenda, and seminar and conference participants at Philadelphia, Milan, New York, FMA, and MME. GAČR grant (403/11/0020, Hanousek) and Global Development Network grant (Kopřiva) support are gratefully acknowledged. The views expressed are those of the authors and do not necessarily reflect the position of any of the affiliated institutions.

[1] While there is a broad consensus that trading on knowledge of, for example, company profits or disclosures is considered insider trading, there is not a similar consensus for trading connected with the execution of large orders or the dual trading practices of some brokers or market makers.

fundamentals (including information on mergers and acquisitions) and information from order flows. PIN may also be affected by large institutional orders, as their presence may have a substantial impact on market microstructure and the price of the asset, particularly on a small emerging market like the Czech Republic. Overall, the extent of information-driven trading considerably affects the credibility of a given financial market, as it also increases the cost of acquiring information on the appropriate timing of a trade.

In the Easley et al. (1996) framework, informed traders act non-strategically and trade upon their inside information. However, informed traders often try to hide their information and react dynamically to the behavior of other market participants, naturally preferring a trading environment with a high degree of anonymity (see Barclay et al., 2003; Anand et al., 2005; Boehmer, 2005; Lee and Yi, 2001; and Brunnermeier and Pederssen, 2005, among others). Hence, an electronic dealers market is an ideal platform for executing informed trades (see also Sherwood, 1997).

Obviously, trades using private information would be not negligible in size. Let us reiterate that the associated PIN is not necessarily an insider measure, as it could also reflect large institutional orders. Typically, on a dealers market large institutional traders cannot hide their orders and as a result would cooperate with a chosen dealer, therefore sharing information about the total order limits (volume and price). Let us note that the execution of such orders is highly dependent upon the particular market microstructure. The current literature does not identify the possibility of collusion between market makers and informed traders. Nevertheless, the particular market microstructure may stimulate investors who are executing large trades to share their private information with a particular market maker (MM). This is the case in the Czech Republic.

In the present paper we use the dealers market (SPAD) of the leading segment of the Prague Stock Exchange, since we believe that the market microstructure of the SPAD trading system might induce collusion between dealers and large institutional investors. In particular, MMs and large investors would share private information and therefore the order flows coming from a given MM may become a significant source and determinant of information-driven trading on small emerging markets.[2]

The Czech capital market microstructure allows investors to place limit orders only on a dealers market, whose trading lots are typically of a small size. Further, as the whole market is quite thin, any large order has a significant impact on the price of the underlying asset. Obviously, executing a large trade through market orders by hitting the quotes of MMs would end up with immense trading costs, as even a few consecutive orders in the same direction would substantially affect the price. While in practice the use of private information could proceed on several fronts, large institutional orders are probably done via one trading channel, i.e., by using one MM. The fast use of private information may lead to a situation where several MMs are informed, investors do not behave strategically, and information is quickly captured by the market. On the other hand, large institutional orders could lead to strategic behavior by the MM, especially on a small market or when liquidity is not large enough.

[2] For example there are publicly known cases where the government of the Czech Republic was selling shares of the energy company CEZ. During a process such as this, one can assume that either the government or the company itself was participating in buying these shares back to keep prices high.

To the best of our knowledge, our study is the first to analyze the extent of order-flow information-driven trading initiated at the level of MMs. We develop a methodology based on the Easley et al. (1996) model to be able to detect suspicious trading behavior by particular MMs on the Prague Stock Exchange (PSE). By an innovative combination of PIN measurement and a jackknife approach we leave out the trades of one particular MM at a time from the sum of all buys and sells. We then test the hypothesis that due to private information about a large block order the MM behaved significantly different from the other MMs, using the estimates from the jackknife approach. Finding significant differences in the behavior of Czech MMs we conclude that contrary to previous studies MMs may not only screen out large informed traders but on less regulated emerging markets they greatly affect the extent of information-driven trading coming from order flows by sharing private information with key large customers. Therefore, our methodology could significantly contribute to the detection mechanisms of order-flow patterns which could be used by other investors as well as regulatory authorities. Our results also contribute to the current debate about market microstructure and its effect on large and small investors.

2. Literature Review

Whenever we talk about informed investors we should distinguish two cases: 1. investors possessing private information originating in firm fundamentals and 2. investors (brokers) accessing information about large institutional orders. Both cases lead to an increase in order flow imbalance, but as mentioned above the second case would probably involve more strategic behavior of the MM. In addition, the second case is interesting to study in the environment of small stock markets, since it is typically associated with dual trading, information advantages that could last for a longer period, possible stealth trading, etc. Below we will present an overview of the relevant literature on information-driven trading, order flows, and stealth trading associated with the behavior of dealers or MMs.

The first stream of literature deals with the problem of whether dual traders are informed or not and how they proceed with large orders. Most of the theoretical studies start with the assumption that dual traders are informed traders and then investigate the effect of their trading strategies (see Roell, 1990, and Sarkar, 1995, among others). Empirical results for developed markets are inconclusive; for example, Fishman and Longstaff (1992) viewed dual trading brokers at the Chicago Board of Trade as informed, while Chakravarty and Li (2003), when controlling for the overall trading profit, suggested that dual traders are uninformed.[3] Nevertheless, an overall view of the literature suggests that MMs or dealers might anticipate private information from the order flow. In addition, informed traders might achieve a more favorable price by breaking up their large orders into multiple medium-sized trades (a so-called "stealth trading" practice; see Barclay and Warner, 1993, for the first reference).[4] The results, however, may vary across different market microstructures; for example, in a pure limit order market (the Stock Exchange of Thailand) informed traders use

[3] The difference between these studies could be associated with the different level of regulation: the earlier study is based on data over a period just before the FBI launched a federal investigation into fraudulent trading practices on the Chicago futures exchange.

[4] For more recent results see Anand and Chakravarty (2007) and Anand et al. (2005), among others.

larger trades compared to dealership markets (Charoenwong, Ding, and Jenwittayaroje, 2010).

The second stream of literature focuses on the overall information advantage of MMs, dealers or brokers rather than on a particular behavior such as dual trading. It is well known that MMs significantly facilitate price discovery compared to a pure auction with only public orders and that their informational advantage comes primarily from the obtained order flow (e.g. Madhavan and Panchapagesan, 2000, and Kurov and Lasser, 2004). Typically, the specialists are able to generate short-term trade profits, mostly as a consequence of the bid-ask spread. Nevertheless, in some markets large dealers act more as informed traders than as liquidity suppliers (see, for example, Wang and Chae, 2003, a study on the Taiwan Stock Exchange). Since only brokers on the market are able to view the order flow of their customers, the informational advantage of the dealers on the market probably originates from the privileged position of direct access to the electronic exchange without any trading fees or trading delays.

Another stream of literature is devoted to the degree of anonymity on different markets and the associated extent of PIN. For example, comparisons of trades on the NYSE and NASDAQ suggest that the NYSE, as a less anonymous market, has a lower extent of informed trading (Garfinkel and Nimalendran, 2003). Moreover, the change in listing from a dealership to an auction market (NASDAQ to NYSE or AMEX) leads to a significant decrease in the extent of information-driven trading. Therefore, either specialists on the NYSE have a better ability to identify informed traders, or informed investors prefer to trade on a market with a higher degree of anonymity (Heidl and Huang, 2002). Similar results were obtained by Grammig et al. (2001) from the Frankfurt Stock Exchange via a comparison of non-anonymous floor trading versus anonymous electronic trading systems (IBIS and later XETRA), showing that informed traders prefer to execute their orders in the anonymous environment.

All of the above-mentioned studies assume that MMs are either using information from the order flow to act against their customers or screening out informed traders. In addition, the results of Hanousek and Podpiera (2002, 2004) support the hypothesis that MMs in an emerging market (the PSE) may share private information with their key large customers. Furthermore, Hanousek and Podpiera (2004) presents more intriguing results – despite many improvements in regulation and increased trading volume the extent of information-driven trading was nearly the same for the years 1999 and 2002. They particularly point out that the extent of informed trading was about the same for shares of Ceska sporitelna and Erste Bank.[5] Let us note that these stocks have little in common except having the same set of MMs, therefore, one could ask to what extent the MMs on the PSE affect the probability of informed trading.

The studies reviewed above suggest that informed traders' behavior differs according to market microstructure and also that MMs are important participants on the market who are able to recognize informed traders. Several studies demonstrate the ability of MMs to identify informed traders and the effect this has on the proba-

[5] In 2000, Ceska sporitelna (a major Czech bank) was privatized to the Austrian Erste Bank. Erste Bank, already listed in Vienna, started dual listing on the PSE in October 2002.

bility of information-driven trading. They conclude that a higher degree of anonymity is associated with a higher probability of information-driven trading, and that informed and insider trading is a widespread practice in emerging financial markets.

3. Methodology

3.1 Easley et al. (1996) Model

Our model is based on the well-known framework developed by Easley et al. (1996). Let us first shortly review their model and then introduce our extension. In all steps of our model, as well as any empirical estimation, we control for the order flow size by assuming/using a regular lot as a trading unit.

There exist three types of agents on the market: uninformed (noisy) traders, informed traders, and MMs. Trading is divided into *n* separate trading days. See *Figure 1* for a tree diagram of the trading day.

Before each day an information event might occur. An information event is defined as the occurrence of a signal *s* about the value of the asset. The probability that a signal occurs is α, and if a signal occurs, it takes on two possible values: low with probability δ and high with probability 1- δ.[6] If a signal occurs, some fraction of the traders receive the signal. If no signal occurs, all traders stay uninformed.

Using the scheme of *Figure 1* we can express the probability of observing a given number of buys and sells as

$$\begin{aligned} L\left((B,S)\,|\,\theta\right) = {} & (1-\alpha) * e^{-\varepsilon T}\frac{(\varepsilon T)^B}{B!}e^{-\varepsilon T}\frac{(\varepsilon T)^S}{S!} && (\textit{no event day}) \\ & + \alpha\delta * e^{-\varepsilon T}\frac{(\varepsilon T)^B}{B!}e^{-(\mu+\varepsilon)T}\frac{((\mu+\varepsilon)T)^S}{S!} && (\textit{bad event day}) \\ & + \alpha(1-\delta) * e^{-(\mu+\varepsilon)T}\frac{\left((\mu+\varepsilon)T\right)^B}{B!}e^{-\varepsilon T}\frac{(\varepsilon T)^S}{S!} && (\textit{good event day}) \end{aligned} \tag{1}$$

where *S* is the number of sells and *B* the number of buys. The first part of expression (1) denotes a no event day, the second part a bad event day and the third part a good event day. According to the assumptions of the model the days are independent and therefore the probability of observing a series of days with a given sum of buys and sells for each day is a product of the probability for the individual days.

$$L\left(B_1,S_1,\ldots,B_I,S_I\,|\,\theta\right) = \prod_{i=1}^{I} L\left((B_i,S_i)|\,\theta\right) \tag{2}$$

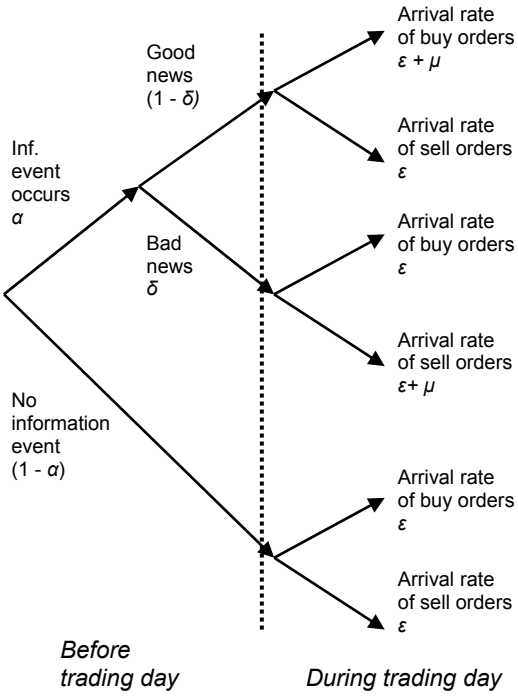
The parameter $\theta = (\alpha, \delta, \varepsilon, \mu)$ is then estimated using the maximum likelihood method.[7] The probability of information-driven trading is the chance that an MM will trade with the informed trader and therefore can be computed as the ratio of the arrival rate of informed traders to the arrival rate of all traders:

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

[6] In the case of a bad signal the value of the asset is $\underline{V}$, for a good signal $\overline{V}$ and for no signal unchanged.

[7] For the estimation we used a rearranged log likelihood function as presented in Easley et al. (2010).

Figure 1 Trading Tree Diagram



Note: This diagram depicts the structure of arriving buy and sell orders during a trading day, where α is the probability of the information event occurring, δ is the probability of bad news, μ is the arrival rate of informed traders and ε is the arrival rate of uninformed traders.

This is actually the conditional probability of an information-driven trade given the occurrence of a trade at the beginning of a trading day. Therefore, the numerator is the product of the probability of an information event times the arrival rate of informed traders. The denominator is then the probability of the occurrence of a trade, which is the probability of an incoming informed trader plus the probability of an incoming uninformed buyer and seller.

3.2 Large Block Trades and Informed MMs

Our extension of the original model reflects the characteristics of a quote-driven market with a relatively small number of market makers. These market makers also usually act as brokers. As on every market there are also various types of investors. We roughly divide them into two groups: large (institutional) investors and small (retail) investors. As these types of investors often have different needs, the market makers (brokers) are specialized on various types of investors or have at least a different approach to these investors. Small investors often open an account at just one broker and either use online trading systems to execute their orders by hitting the quotes posted by market makers or use some of the brokerage services. Nevertheless, large investors have specific needs in order to execute their large orders, as these orders often have a significant price impact.

Large investors are therefore insiders, as they possess valuable private information (coming from the order flow), which has a significant effect on the price of the asset (Chakravarty, 2001, and Golec, 2007, among others). According to their size, they are the key customers for the brokers. Therefore, they may get special brokerage services and the fees for executing orders may differ significantly from those charged to retail customers. According to Schwartz and Shapiro (1992), institutional investors accounted for 72% of the share volume on the New York Stock Exchange. According to the relatively low number of retail investors in the Czech Republic, this percentage might be significantly higher.

At some markets the problem of the different needs of different types of investors and the potential problem of better-informed large investors with a large impact on prices is solved by an upstairs market (Golec, 2007). However, similar to the market microstructure of the Czech capital market, in our model there is no upstairs market for large orders. Similar to the existing literature on stealth trading, we expect that large orders are usually broken into medium-size trades. As large investors face a decision how to optimize the execution costs of their orders, using market orders (similar to small investors) to execute large trades is not usually suitable for them due to the large impact on the price and therefore the large execution costs. Thus, large investors may often prefer passive trading strategies, i.e., using limit orders. Nevertheless, on the quote-driven market only market makers may place limit orders (quotes). Therefore, large investors are forced to seek lower execution costs by negotiating with market makers. In other words, while optimizing the execution costs of their orders, large investors are in our model forced to negotiate with market makers about the possible execution of the orders and market makers may also use market making activities in order to execute these orders. That is, there is latitude for collusion or large investors could be forced to cooperate and share their information with the MM. Some of these facts are also mentioned, for example, by Keim and Madhavan (1995), Chan and Lakonishok (1995), and Golec (2007), confirming that larger trades take several days to execute.

Such implementation of large trade orders has a higher chance to minimize the impact on the stock price; practically it means that the MM trades against his account and once he secures the deal (accumulates or sells shares) then a block trade with his client closes the trade. As this scenario is in fact using the market maker's activities to hide the large trade, it is not rational for the large investor to contact more than one market maker/broker, as it would just spread the information about an incoming large order. Therefore, at the beginning of the execution of his large order, the large investor chooses a trading channel, meaning he chooses a market maker, which we can distinguish in our dataset. In other words, the MM has an incentive to act strategically in that he is trying to choose the optimal timing of several trades to process the whole big order at the best possible price.[8]

Our model, therefore, consists of two types of private information. The first one is the private signal in the Easley et al. (1996) model: short-lived information about the underlying asset that is available to a relatively large number of informed

[8] However, we do not expect that he is necessarily trying to manipulate the price or abusing the market illegally.

investors. Similar to Keim and Madhavan (1995), we assume that in such a situation investors prefer market orders and execute their trades as quickly as possible. In other words they do not act strategically and come to the market according to the private information they received. The second type of private information in our model comes from the order flow. Information about the incoming large order is available only to the large investor and the market maker with whom the large investor is negotiating about the optimal execution of the trade. Similar to Keim and Madhavan (1995), we assume that an investor who possesses longer-term private information that is available only to a restricted number of people prefers to trade more discreetly, negotiating with some of the market makers on the preferable execution of the trade. Also, large orders executed through market orders have large price impacts and, therefore, using limit orders may significantly reduce the execution costs of the trade. Keim and Madahavan (1995) argue that the benefits from a passive trading strategy (limit orders) should be largest on thin markets where liquidity is low. Their analysis of the data on the equity transactions of 21 institutions shows that the execution time of trades is longer than one day: on average 1.65 to 1.80 days. However, this might be significantly higher for the Czech Republic. As pointed out above, the institutional investor often just announces the number of shares and the side of the trade and the broker then tries to execute the order with the lowest possible execution costs.

The first type of information flows to the market through market orders and therefore its revelation is not affected by the behavior of particular market makers. On the other hand, the large investor, when optimizing his execution cost, chooses one of the market makers/brokers, who is then executing his order and therefore has other incentives than to balance his portfolio. In such a situation we in fact have two types of MMs: informed and uninformed.

Suppose that there is other information affecting the price of an asset: information about a large order that is independent of the above private signal of informed investors and that lasts for several trading days. In such a situation the large investor will contact just one MM, as otherwise he would be spreading the information to the other participants of the market, which could increase the execution costs of the trade. Therefore, we assume that only one informed MM has private information about this large order coming on the market from one of his clients. The large order consists of a random volume of shares and a random length K of trading days. Note that the actual number of days and the total number of shares can be limited by the price ceiling imposed by the client or/and by a particular deadline.[9] As confirmed by several brokers, the typical practice is that a large order is inspected at the end of the trading session and new limits are set for the next day(s) or the execution of the order is stopped. In such a situation only one MM will have detailed information about the large order – private information. If more than one MM receives information about the large order and if the MMs do not act in consonance with each other, the order will be revealed to the whole market and the new value of the asset will be revealed immediately by the competitive behavior of two or more informed MMs.[10]

[9] Even though block trades must be reported in 5 minutes in the open session and in 60 minutes in the closed session, the behavior of MMs suggests that they are either aware of the block trade in advance or set the block trade *ex-post*.

If the MM is informed, we assume that he does not set quotes in a way that will immediately reveal his information about the order. Therefore, in the case of a large buy order the informed MM will just try to have the best quote[11] – he will post his quotes for buys more actively and thereby end up with the best quote with a higher probability than the uninformed MM. Although the other MMs may anticipate the existence of the large order, they do not know the exact information of the trade, i.e., the limit price and the execution deadline. This is the key information for the other market participants to be able to use such information to actively post quotes for buys and compete with the informed MM. Without this information, there is the risk that the MM will immediately stop the execution of the trade and the price will return to its initial level.

Further, the uninformed MMs might also find it difficult to compete with the price setting of the informed market maker, as they do not possess the inventory advantage of the informed MM. The uninformed MMs without inside information about the large order will try to avoid risky unbalanced positions and therefore will post quotes such that they would finish with somewhat balanced inventories. On the other hand, the informed MMs, contingent on their information, might venture more risky positions from the point of view of uninformed MMs and therefore might be able to afford to actively quote only buys or sells. The uninformed MMs generate profit from the trading fees and spread. However, the informed MM generates profit also from proprietary trading. Thus he may force the other MMs either to accept a lower spread and thus lower trading profits or to give up market making activities for the particular stock. For the other investors it is also quite hard to trade upon just part of the information – the information that one of the market makers might be executing a large order – as they would face the spread costs while trading with only this incomplete information. They could, however, use such information to postpone the execution of their trade.

Therefore, such a situation will probably end up such that the informed MM uses his market making activities to execute the trade and thus will probably have a different balance of mandatory buys and sells than the rest of the market makers. Although the other market participants might be aware of the presence of a large order it could be difficult to use this information in the current trading system. Since the uninformed market makers have limited resources and therefore have to balance their inventory position it is really hard for them to compete with the informed market maker. If an uninformed MM would like to trade using this information he would face the risk that the large investor will stop selling the share. Here it should be stressed again that, as confirmed by several brokers, the typical practice is that the large order is inspected at the end of the trading session and new limits are set for the next day(s). It is this – the limited access to the limit orders of all investors, uninformed market makers with limited capital who raise money mostly from the spread, the non-existent upstairs market, and the possibility of non-transparent pseudo block trades that move from block trades into the SPAD segment – that makes the market

[10] Given the trading environment (a dealers' market) we expect a relatively low number of market makers, therefore, due to the competition of two or more informed market makers, we expect that the information about the large order will be revealed quite quickly.

[11] The best buy and best sell quotes from all the market maker's quotes.

Table 1 Total Traded Volume and Mandatory Trades, Czech Government Selling Shares of CEZ

Time Span	Total Volume (bil. CZK)	Mandatory Trades
		76.6
Sep 07–Aug 08	373.9	148.8
		72.2

Note: In the Mandatory Trades column, the main number is the overall volume while the number in the upper right corner is the mandatory buy volume and the number in the lower right corner is the mandatory sell volume.

Source: www.akcie.cz and authors' computations

microstructure really messy and enables some market makers to use their market making activities to cover the information stemming from large trades.

Overall, since the major source of the income of an MM is generated by the spread between the bid and ask prices, the situation of an existing large trade conducted via a specific MM could lead to 1. collusion, 2. the other MMs' spreads being close to their marginal costs, and/or 3. other MMs stopping trading. The informed MM, on the other hand, will trade actively only on one side (buy or sell), according to his private information. By and large it leads to a market with lower competition and large latitude for price manipulation.

To illustrate the situation let us focus on the particular example of a publicly known case of the Czech government selling shares of the energy company CEZ. The publicly available information was that the Czech government started selling nearly 7% of the shares in September 2007 and stopped selling the shares due to the upcoming financial crises by September 2008. One of the commissioners helping the Czech government with the execution of the trade was MM4.[12] The overall volume of the trade was around CZK 34 billion, i.e., 9% of the total volume traded during September 2007 to September 2008 (see *Table 1*). By comparing the trading volumes by category one can speculate that only a part of the governmental deal was conducted through mandatory trades; most of trades during this period probably used the MM4 quotes as an indication of the limit orders of the Czech government.

Further, the fact that during this period the percentage of the market share of MM4 increased to 18% (overall 13% for CEZ) supports our hypothesis that the inventory advantage and the detailed private information about large orders give the informed MM the opportunity to trade more actively and end up with an unbalanced inventory position (see *Table 2*). The execution of the trade ended with the first large price jump in September 2008, which also supports the notion that a large order is inspected at the end of the trading session and stopped when the price changes significantly. If any uninformed MM would like to trade using the publicly available information that the Czech government is going to sell the shares he would face the risk that MM4 will stop selling the share. This is probably the case we see in the behavior of the market makers as the governmental deal was conducted, since the price remained constant for nearly the whole period. MM4 appeared significantly

[12] We decided to code all the market makers in order to minimize possible bias while executing all computations and analysis. The coding of market makers is available upon request.

Table 2 Mandatory Trades of Particular Mms, Czech Government Selling Shares of CEZ

MM	MM1	MM2	MM4	MM5	MM6	MM7	MM9	MM10	MM11
From Sep 07 to Aug 08	9.7	6.4	18.3	5.5	8.5	11.1	6.0	7.0	4.2
	1.8	0.6	10.4	0.1	-0.8	-3.8	0.5	-4.5	0.0
	7.8	5.8	7.9	5.3	9.3	14.9	5.5	11.5	4.2

Note: In each column the main number represents the difference between mandatory buy and mandatory sell volume while the number in the upper right corner is the mandatory buy volume and the number in the lower right corner is the mandatory sell volume.

Source: www.akcie.cz and authors' computations

more on the quotes for buys (market maker selling the asset) and MM7 and MM10 (large market makers) appeared mostly on the quotes for sells (market maker buying the asset).

3.3 Estimation Procedure

To estimate the extent of information-driven trading due to large orders or, in other words, due to informed MMs, first we run the estimations for the whole sum of buys and sells. Further, to estimate the PIN originating from large orders or other private information of the MMs we propose a procedure to estimate the PIN with and without the trades of informed MMs. Therefore, step by step we exclude each MM's trades from the sum of buys and sells and estimate the model. Having all the parameters $\theta_i = (\alpha_i, \delta_i, \varepsilon_i, \mu_i)$ estimated for each MM we then test whether PIN using the estimated parameters $\theta = (\alpha, \delta, \varepsilon, \mu)$ and PIN without considering the trades of a given MM are significantly different.

Having identified the informed MMs we can estimate the effect of the large orders on the probability of information-driven trading:

$$PIMM = \left| \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} - \frac{\alpha_i\mu_i}{\alpha_i\mu_i + 2\varepsilon_i} \right| \tag{4}$$

where $\theta = (\alpha, \delta, \varepsilon, \mu)$ are the estimated parameters from the classic Easley et al. (1996) model using the sum of all buys and sells for each day and $\theta_i = (\alpha_i, \delta_i, \varepsilon_i, \mu_i)$ are the estimated parameters using the sum of all buys and sells for each day without the trades of a given identified informed MM. We have made the same set of assumptions as Easley at al. (1996); in particular, we assume independence of information arrival(s). Therefore, both parts of (4), i.e., $\frac{\alpha\mu}{\alpha\mu + 2\varepsilon}$ and $\frac{\alpha_i\mu_i}{\alpha_i\mu_i + 2\varepsilon_i}$ would have asymptotically normal distributions and obviously the estimators of both parts would be (positively) correlated.[13] In order to get the standard error of (4) we use a cluster modification of the sandwich estimator as proposed by Rogers (1993) to estimate the joint covariance matrix of $\frac{\alpha\mu}{\alpha\mu + 2\varepsilon}$ and $\frac{\alpha_i\mu_i}{\alpha_i\mu_i + 2\varepsilon_i}$. Non-technically speaking, the extent of information-driven trading coming from the behavior of an informed

MM (equation 4) is therefore the difference between the probability of informed trading with and without the trades of the informed MM.

4. Data

For our analysis, we use intra-day data from the Prague Stock Exchange (PSE) SPAD trading system for all stocks traded from January 1, 2003 to August 31, 2010, publicly available online.[14] SPAD was founded in 1998 to increase the liquidity of the market. The trading system is designed as a dealers market with at least three MMs for each stock, who are required to quote ask and bid prices for a standardized number of shares with a limited maximum possible spread for each stock. If a given quote is the best available on the market the particular MM is obliged to trade on the posted quote for a buy or sell.

Each trading day is divided into two phases, open and closed. The actual trading occurs during the open phase of the system, from 9.15 a.m. to 4.00 p.m. each trading day. We use data on all SPAD trades during the sample time period. Each trade record in our database consists of security identification, date, time, type of trade, and price, and for the standard SPAD trades also the identification of the MM who traded it. We are also able to identify cross trades and trades conducted between the inventory of the MM and the MM's client. Even though akcie.cz provides quite detailed information about trades, still a significant proportion are not transparent, as mandatory trades account for only around 40% of the total traded volume. Despite the increased regulation of block trades these trades probably just moved from the segment of block trades into the non-transparent segment of SPAD trades with no identification. The key feature of our dataset is that we are able to identify not only whether the given trade was buyer- or seller-initiated, but also which MM was on which side of the trade.

The sample period consists of 1925 trading days and we focus on 15 companies traded during the period (see *Table 3a* and *3b* for descriptive statistics of market capitalization and traded volumes; KIT DIGITAL was traded only for a short period in our dataset).[15] We have eleven MMs in our sample period: six brokerage

[13] One can get this statement implicitly from previous studies, as for testing PIN all authors used *t*-ratios and underlying *t*-tests. Also, the assumption of independence of information arrival ensures that the first and second terms of (4) represent certain conditional probabilities; for example, the first term is actually the conditional probability of an information-driven trade given the occurrence of a trade at the beginning of a trading day. Therefore, such conditional probabilities can also be computed as the means of some number of buys or sells satisfying certain conditions. Since we assume independence of buy and sell orders, any (artificial) mean using functions of the number of buy and sell orders would satisfy the CLT and therefore we would get asymptotic normality of PIN.

[14] Available at www.akcie.cz. The last access for this paper was on 30 September 2010.

[15] Let us note that only six of them were traded during the whole period: two banks (Erste Bank and Komercni banka), a petrochemical company (Unipetrol), an electricity producer (CEZ), a telecommunications company (Telefonica O_2), and a cigarette producer (Philip Morris). Another telecommunications company (Ceske Radiokomunikace) was removed from the market in September 2004. One IPO, Zentiva, was introduced to the market in June 2004 and removed from the market in April 2009. In February 2005 a real estate company (ORCO), already listed in Paris, started dual listing on the PSE and in June 2005 a media company (CME), already traded on NASDAQ for over 10 years, started dual trading on the PSE. ECM (a real estate company) and PEGAS (a synthetic nonwoven textiles producer) were introduced in December 2006, AAA (a car reseller) in September 2007, NWR (a coal mining company) in 2008, and VIG (an insurance company) in 2008.

Table 3a Market Capitalization and Overall Traded Volumes

Stock	Year	Mkt. cap.	Turn-over	SPAD trades	Sys. trades	APD	B/S	Price	Price change	MM
AAA	2007	3	22%	0.93	0.34	14.4	0.89	44.4	na	7
	2008	1	75%	0.94	0.61	13.2	0.82	9.1	-80%	7
	2009	1	13%	1.00	0.90	7.8	0.92	13.5	49%	6
CME	2005	43	14%	0.81	0.56	18.3	1.30	1,409	18%	6
	2006	50	50%	0.96	0.64	42.2	1.05	1,462	4%	6
	2007	72	44%	0.99	0.68	44.3	1.12	2,106	44%	6
	2008	14	130%	0.98	0.52	29.0	0.85	382	-82%	6
	2009	25	69%	0.99	0.60	71.0	0.99	446	17%	6
CEZ	2003	86	51%	0.64	0.31	18.4	1.10	146	58%	10
	2004	202	54%	0.73	0.42	42.0	1.30	341	134%	9
	2005	436	69%	0.68	0.45	124.5	1.00	736	116%	10
	2006	569	61%	0.94	0.48	157.8	0.96	960	30%	9
	2007	807	50%	0.98	0.44	119.1	1.02	1,362	42%	9
	2008	465	83%	0.99	0.44	110.9	0.93	785	-42%	9
	2009	465	44%	1.00	0.41	66.5	0.99	864	10%	9
CRA	2003	11	45%	0.72	0.25	5.3	2.00	345	83%	8
	2004	14	67%	0.61	0.29	9.8	1.10	444	29%	8
ECM	2006	5	50%	0.99	0.53	84.4	1.51	1,432	na	6
	2007	5	351%	0.99	0.60	44.7	0.95	1,203	-16%	6
	2008	2	285%	0.98	0.51	25.0	0.61	261	-78%	7
	2009	2	26%	0.98	0.80	8.2	1.05	308	18%	7
EB	2003	191	7%	0.78	0.61	17.4	1.20	798	59%	6
	2004	287	11%	0.85	0.63	31.6	1.20	1,187	49%	6
	2005	334	14%	0.83	0.63	43.5	1.00	1,372	16%	8
	2006	505	12%	0.94	0.66	54.0	1.05	1,601	17%	9
	2007	411	25%	0.99	0.71	92.5	0.94	1,301	-19%	9
	2008	133	68%	1.00	0.65	114.6	0.94	419	-68%	9
	2009	264	22%	0.99	0.62	100.4	1.01	699	67%	9
KB	2003	92	110%	0.65	0.40	38.0	1.00	2,418	16%	9
	2004	124	120%	0.60	0.34	61.1	1.00	3,272	35%	9
	2005	131	158%	0.64	0.43	95.1	0.90	3,441	5%	10
	2006	118	90%	0.94	0.57	68.8	0.92	3,099	-10%	9
	2007	166	82%	0.99	0.57	76.2	1.13	4,371	41%	8
	2008	113	102%	0.99	0.59	100.7	0.97	2,970	-32%	7
	2009	149	42%	0.99	0.50	74.3	1.06	3,929	32%	6

Notes: Mkt. cap. is Market capitalization in billions of CZK at the end of the year; *Turnover* is the turnover ratio during the year as a percentage of Mkt. cap.; *SPAD trades* is the ratio of the SPAD traded volume on overall traded volume; *Sys. Trades* is the ratio of system trades (usually classic trades with an identification of the market maker) to the overall traded volume; *APD* is the average number of trades during a trading day; *B/S* is the buy over sells ratio; *Price* is the price at the end of the year; *Price change* is the percentage change of the price during the year; *MM* is the id number of the MM.

Sources: PSE fact books, www.akcie.cz and authors' computations

Table 3b Market Capitalization and Overall Traded Volumes

Stock	Year	Mkt. cap.	Turn-over	SPAD trades	Sys. trades	APD	B/S	Price	Price change	MM
NWR	2008	19	229%	1.00	0.59	101.1	0.90	73	na	7
	2009	43	64%	0.99	0.65	87.2	1.02	162	120%	7
O_2	2003	94	69%	0.49	0.17	22.5	1.20	291	19%	10
	2004	119	102%	0.52	0.16	35.9	1.20	369	27%	9
	2005	169	171%	0.44	0.14	43.0	1.00	525	42%	10
	2006	153	64%	0.88	0.41	62.8	0.92	476	-9%	10
	2007	175	58%	0.99	0.40	51.4	1.13	545	14%	10
	2008	137	68%	0.99	0.38	52.3	0.91	424	-22%	10
	2009	135	49%	0.99	0.40	38.6	1.01	418	-1%	10
ORCO	2005	na	na	0.78	0.61	18.8	1.10	1,809	41%	6
	2006	22	125%	0.96	0.76	65.1	1.09	2,755	52%	6
	2007	23	159%	0.99	0.64	54.9	0.87	2,165	-21%	8
	2008	2	630%	1.00	0.64	52.5	0.79	173	-92%	8
	2009	2	109%	1.00	0.74	34.0	1.06	170	-1%	8
PM	2003	30	64%	0.67	0.38	9.1	1.20	15,728	41%	9
	2004	32	91%	0.72	0.41	22.2	1.10	16,776	7%	8
	2005	35	101%	0.68	0.43	28.2	1.20	18,251	9%	8
	2006	21	89%	0.93	0.46	23.4	0.92	10,840	-41%	7
	2007	15	62%	0.99	0.47	16.0	1.08	7,933	-27%	7
	2008	12	42%	0.94	0.38	10.1	0.92	6,026	-24%	7
	2009	17	30%	0.94	0.39	9.1	1.03	8,796	46%	7
PEGAS	2006	7	48%	0.99	0.35	93.3	0.91	753	na	7
	2007	7	231%	0.95	0.48	35.0	0.93	751	0%	8
	2008	2	237%	0.97	0.57	21.9	0.78	233	-69%	8
	2009	4	71%	0.95	0.51	12.0	1.19	445	91%	8
UNI	2003	12	72%	0.60	0.34	8.2	1.30	66	92%	8
	2004	18	79%	0.68	0.35	9.5	1.00	98	48%	8
	2005	42	122%	0.78	0.54	45.1	1.00	233	138%	8
	2006	42	114%	0.95	0.59	47.2	1.01	234	0%	6
	2007	61	75%	0.88	0.55	33.3	1.05	338	44%	7
	2008	27	105%	0.98	0.60	28.0	0.90	150	-56%	7
	2009	25	54%	0.98	0.53	17.7	1.11	140	-7%	7
VIG	2008	83	2%	1.00	0.76	11.6	0.84	646	na	5
	2009	121	1%	1.00	0.79	8.3	0.98	942	46%	5
ZEN	2004	29	59%	0.65	0.30	17.0	1.10	758	50%	8
	2005	43	232%	0.61	0.38	48.5	1.00	1,136	50%	8
	2006	48	222%	0.94	0.60	71.7	1.01	1,268	12%	9
	2007	37	300%	0.99	0.54	71.5	1.00	972	-23%	9
	2008	41	116%	0.92	0.33	28.5	1.09	1,078	11%	9

Notes and *Sources* – see *Table 3a*.

Table 4 Market Share of Market Makers on the PSE During the Sample Period

Stock	AAA	CME	CEZ	CRA	ECM	EB	KB	O2	ORCO	NWR	PM	PN	UNI	VIG	ZEN
MM 1	**13%**	**24%**	**13%**	**12%**	**19%**	**16%**	**17%**	**11%**	**19%**	**20%**	**14%**	**17%**	**19%**	**28%**	**15%**
	599	14544	21891	297	3517	20258	24664	8952	9763	9262	4478	2750	9612	1462	8726
MM 2	**6%**	**12%**	**9%**	**8%**	**0%**	**9%**	**2%**	**7%**	**2%**	**3%**	**10%**	**11%**	**11%**		**8%**
	222	6271	15131	199	114	12052	1985	5948	1070	1883	3389	2057	5382		4903
MM 3			**0%**	**2%**			**0%**	**0%**			**1%**		**0%**		
			205	46			422	364			113		112		
MM 4	**9%**	**14%**	**13%**	**14%**	**17%**	**12%**	**14%**	**13%**	**16%**	**13%**	**13%**	**14%**	**16%**	**19%**	**14%**
	555	7745	20703	355	3455	15005	20066	10307	8687	6285	4248	2514	7813	1092	8064
MM 5	**26%**	**15%**	**9%**	**14%**	**15%**	**15%**	**12%**	**11%**	**17%**	**19%**	**13%**	**12%**	**18%**	**15%**	**9%**
	2447	10821	16578	348	4220	20477	17868	9607	12779	11915	4308	2483	9215	848	5618
MM 6	**10%**	**0%**	**10%**	**7%**	**17%**	**7%**	**12%**	**9%**	**2%**	**9%**	**11%**	**6%**	**5%**		**10%**
	483	84	17367	169	2894	9427	17371	7173	1441	3747	3618	1351	2326		5943
MM 7	**30%**	**19%**	**20%**	**16%**	**17%**	**16%**	**16%**	**25%**	**17%**	**25%**	**18%**	**19%**	**16%**	**21%**	**15%**
	2017	12894	32078	423	4136	21734	24112	20021	12033	13958	6263	3711	8606	1207	8878
MM 8			**0%**				**1%**	**3%**							
			885				857	2243							
MM 9			**7%**	**10%**		**7%**	**8%**	**6%**	**12%**	**1%**	**5%**	**10%**	**2%**		**10%**
			12865	258		8537	10280	4722	5125	354	1240	1659	1025		6024
MM 10	**7%**	**14%**	**13%**	**18%**	**15%**	**13%**	**14%**	**13%**	**15%**	**11%**	**14%**		**13%**	**17%**	**13%**
	267	7500	22088	475	2904	15859	20736	10407	7142	5326	4473		6279	931	7908
MM 11			**6%**			**4%**	**4%**	**4%**							**5%**
			9960			4187	4884	3097							2916

Note: Each row consists of the percentage of volume traded through mandatory trades and the number of mandatory trades of a given market maker during the sample period 1 January 2003 to 30 August 2010.

Source: www.akcie.cz and authors' computations

firms and five banks.[16] The MMs also differ in their specialization in different types of customer: retail vs. large institutional investors.

As we can see from *Table 4* none of the eleven MMs on SPAD had a significantly higher market share in any of the analyzed stocks. The maximum market share was about 30 percent for one MM and each traded stock had at least five MMs with a more or less comparable market share.

The average number of trades during a day differs significantly among the stocks during the sample period. Only some of the newly introduced stocks attracted the attention of investors quickly and the activity of some of these new blue chips on the PSE was not comparable to the already established stocks. Our model assumes a significant role of block trades as a source of information for some MMs, and the data seem to confirm this assumption. Block trades are defined by a limit set by the PSE and this limit is considerably larger than the market capitalization of the trading lots in SPAD. According to current regulations every block trade has to be registered within 5 minutes during the open phase (9.15 a.m. to 4.00 p.m.) and within 60 minutes during the closed phase.

Table 5 clearly shows that a significant percentage of the volume traded on SPAD used block trades. One can speculate that during 2003 to 2005, when there was a high percentage, the MMs who were focused on large customers also used standard SPAD trades to be able to gather stocks in order to execute block trades. Such MMs are actually informed traders and thus the block trades may have been an indication of a high level of private information on the PSE. Let us comment that the significant decrease in the percentage of block trades in 2006 was probably caused by the increased regulation of MMs.[17] Despite this increased regulation, the suspicious practices of unidentified block trades are still present. The forbidden unidentified block trades have moved to the segment of the SPAD market with no identification of the trading parties, leading to a situation similar to the one before 2006 (see *Table 5*). Therefore, trading practices probably remained the same, only the placement and reporting of the trades changed.

SPAD was introduced to increase liquidity on the PSE. However, due to the size of the trading lots only medium-sized and large investors could trade in the system. As *Table 6* demonstrates, the trading lots have varied quite a lot, as the prices of some stocks grew significantly during the sample period. For example, the smallest trading lot (AAA) started at CZK 0.01 million, while the largest lot was CZK 7.18 million (CEZ).[18] Such variance in the mandatory minimum trading volume is another problem of the SPAD trading system as it might present obstacles for uninformed and smaller investors; the design of the market then attracts mostly large institutional and informed investors. For this reason the effect of changing the lot size can significantly affect the extent of information-driven trading, as according to the Easley et al.

[16] The brokerage firms are ATLANTIK finanční trhy, a.s., BH Securities a.s., CA IB Securities, a.s., Fio, burzovní společnost, a.s., Patria Finance, a.s., and WOOD & Company Financial Services, a.s.; the banks are Česká spořitelna, a.s., HVB Bank Czech Republic a.s., Raiffeisenbank a.s., ING Bank N.V., and Komerční banka, a.s.

[17] Since early 2006, all MMs and brokers have been obligated to report their activities to the regulation authority, including their dealings book.

[18] Using the average exchange rate against USD over the period studied (~23.7 CZK = 1 USD), the lot size varies from USD 25,000 to USD 232,000.

Table 5 SPAD Traded Volume and Percentage of Block Trades

Stock	Year	Volume bil. CZK	Block trades	SPAD with ID	SPAD no ID	SPAD ID no cross
AAA	07-10	1.1	**6%**	49%	45%	**42%**
CME	2005	5.9	**17%**	55%	26%	**52%**
CME	06-10	98.8	**2%**	62%	36%	**57%**
CEZ	03-05	445.6	**30%**	42%	27%	**40%**
CEZ	06-10	1343.3	**2%**	45%	53%	**42%**
CRA	03-05	14.2	**32%**	25%	41%	**23%**
ECM	06-10	24.0	**1%**	58%	41%	**53%**
EB	03-05	91.4	**16%**	62%	20%	**59%**
EB	06-10	318.7	**1%**	67%	32%	**62%**
KB	03-05	448.6	**37%**	39%	24%	**36%**
KB	06-10	445.9	**2%**	56%	42%	**52%**
NWR	08-10	89.5	**0%**	62%	37%	**54%**
O_2	03-05	472.4	**53%**	14%	32%	**13%**
O_2	06-10	365.1	**4%**	40%	56%	**37%**
ORCO	2005	5.6	**20%**	60%	19%	**56%**
ORCO	06-10	74.6	**2%**	68%	30%	**63%**
PM	03-05	82.7	**30%**	39%	30%	**36%**
PM	06-10	38.9	**5%**	44%	51%	**41%**
PEGAS	06-10	25.3	**4%**	48%	47%	**44%**
UNI	03-05	70.9	**22%**	45%	31%	**41%**
UNI	06-10	133.9	**6%**	56%	37%	**52%**
VIG	08-10	3.3	**0%**	76%	23%	**69%**
ZEN	03-05	119.9	**39%**	35%	25%	**33%**
ZEN	06-09	260.5	**4%**	53%	43%	**50%**

Note: Volume is the traded volume on SPAD; *Block trades* is the percentage of the SPAD volume; *SPAD with ID* (*no ID*) is the percentage of SPAD traded volume with (without) an identification of the market maker; *SPAD ID no cross* is the percentage of SPAD traded volume analyzed in our study (standard SPAD trades through the market maker).

Source: www.akcie.cz and authors' computations.

(1996) model informed traders are more likely to trade larger volumes. Therefore, regarding the significant increase of retail investors in the Czech Republic, lowering the lot size might attract more uninformed investors, as on SPAD the fees are significantly lower compared to other trading channels.

5. Results

Trading at the PSE in general, the structure of potential investors, and the behavior of MMs follow specific and significantly different patterns during the morning and afternoon sessions. Therefore, we decided to estimate the extent of information-driven trading for each session separately. Basically, new information comes to the Czech capital market before the morning session and then again in the afternoon when there is news from U.S. capital markets. Note that only a negligible fraction of trades takes place between 12.00 p.m. and 2.00 p.m., and even these are mainly automatic. Therefore, we decided to divide each day into two main parts, the morning session from 9.15 a.m. to 12.00 p.m. and the afternoon session from 2.00 p.m. to 4.00 p.m., to better reflect trading

Table 6 Changes in the Trading Lot Size

Stock	Time period	LOT size	Price (CZK)		Volume mil. CZK	
			Min	Max	Min	Max
AAA	Sep 07–Aug 10	**3,000**	4.8	56.5	0.01	0.17
CME	Jun 05–Aug 10	**1,000**	104	2317	0.10	2.32
CEZ	Jan 03–Oct 04	**20,000**	87	282	1.74	5.63
	Oct 04–Aug 05	**10,000**	257	552	2.57	5.52
	Aug 05–Aug 10	**5,000**	523	1435	2.61	7.18
CRA	Jan 03–Sep 04	**3,000**	180	535	0.54	1.61
ECM	Dec 06–Aug 10	**500**	120	2065	0.06	1.03
EB	Jan 03–Sep 03	**500**	1850	2975	0.93	1.49
	Sep 03–Mar 04	**1,000**	2685	3886	2.69	3.89
	Mar 04–Jul 04	**500**	3530	4236	1.77	2.12
	Jul 04–Aug 10	**2000***	196.1*	1743*	0.39	3.49
KB	Jan 03–Sep 03	**2,000**	1817	2680	3.63	5.36
	Sep 03–Jun 08	**1,000**	2210	4540	2.21	4.54
	Jun 08–Aug 10	**500**	1520	4295	0.76	2.15
O_2	Jan 03–Aug 10	**5,000**	240	628	1.20	3.14
NWR	May 08–Aug 10	**5,000**	59	624	0.30	3.12
ORCO	Feb 05–Aug 10	**500**	70	3785	0.04	1.89
PM	Jan 03–Mar 04	**200**	10400	20740	2.08	4.15
	Mar 04–Aug 10	**100**	3650	21451	0.37	2.15
PEGAS	Dec 06–Aug 10	**1,000**	166	848	0.17	0.85
UNI	Jan 03–Feb 05	**20,000**	34	181	0.68	3.63
	Feb 05–Aug 10	**10,000**	89	346	0.89	3.46
VIG	Feb 08–Aug 10	**500**	400	1478	0.20	0.74
ZEN	Jun 04–Jun 07	**3,000**	480	1571	1.44	4.71
	Jun 07–Apr 09	**2,000**	784	1448	1.57	2.90

Notes: LOT is the number of shares in the trading lot; *Price* and *Volume Min (Max)* is the minimum (maximum) price and volume in CZK during the corresponding time period.
* indicates stock splitting.

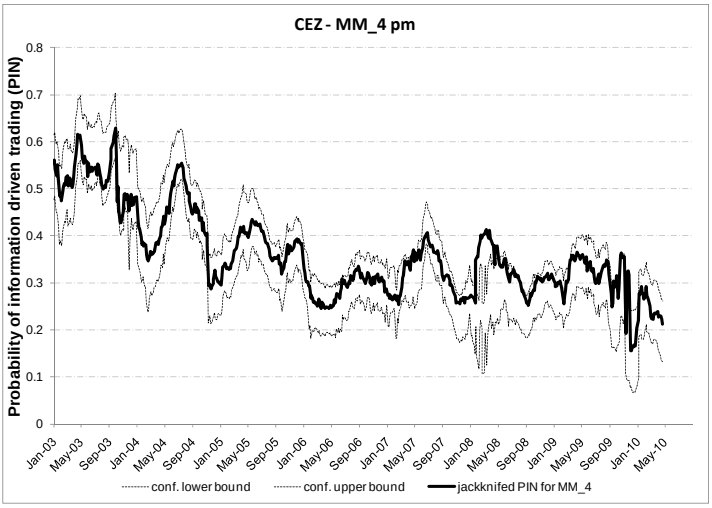
Source: www.akcie.cz and authors' computations

patterns and the specific nature of a small emerging market with a substantial foreign presence.

For (automatic) identification purposes we first run a rolling window of 90 trading days through our sample period and for each window estimate the extent of information-driven trading. We believe that the 90-trading day window is an optimal balance between the assumption of the underlying Poisson process being stationary and the length of the estimation moving window, which affects the precision of the estimates.[19] The results are graphically presented in *Figures 2* and *3*. Based on the patterns visible in these figures, we focus on particular stocks for which the rolling window

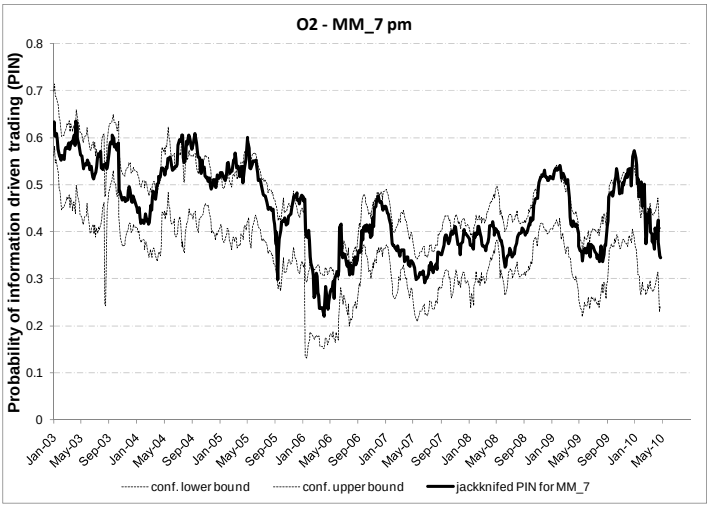
[19] We have run the estimation also for shorter rolling windows, nevertheless our results suggest that that the 90-day rolling window still satisfies the assumptions of the model, as the results are similar for the shorter rolling windows. Detailed results are available upon request.

Figure 2 Estimated PIMM for CEZ Using a 90-Day Rolling Window, Afternoon Trading



Note: This figure represents a graphical version of the test; suspicious behavior is identified when results for particular MMs (thick line) exceed the limits of the confidence interval.

Figure 3 Estimated PIMM for Telefonica O_2 Using a 90-Day Rolling Window, Afternoon Trading



Note: This figure represents a graphical version of the test; suspicious behavior is identified when results for particular MMs (thick line) exceed the limits of the confidence interval.

analysis suggested significantly different behavior of particular MMs. While inspecting these figures, we could observe that for shares for which the PSE is the main market the afternoon sessions always depict higher PIN. Secondly, by the same token, the dual traded shares show just the opposite pattern, i.e., the morning session has higher PIN. Overall, possibly due to the strengthening of the regulation of MMs

by introducing the requirement to regularly report detailed information about their activities, the extent of information-driven trading decreased significantly until 2008, when a huge decrease of traded volume due to the financial crisis again increased the PIN of most of the stocks.

Let us note that our trading data consists of precise information on whether the trade is a mandatory buy or mandatory sell, contrary to most of the existing studies.[20] Boehmer et al. (2007) point out that using only the estimation of whether the trade is buyer- or seller-initiated leads to downward-biased PIN estimates and that the magnitude of the bias is related to the security's trading intensity. This may partly explain the difference in our results compared to the results of Hanousek and Podpiera (2004), as they used data for the whole day and estimated whether the trade was buyer- or seller-initiated using Lee and Ready (1991)'s methodology. Hanousek and Podpiera (2004) concluded that during 1999–2002 they did not see any improvement in the extent of information-driven trading. Nevertheless, our results suggest that all the blue chips experienced a significant decrease in PIN during 2003–2006.

In *Tables 7a* and *7b* we present the results of tests for the time periods and stocks identified in the automatic identification phase described above. Overall, our results suggest that during our sample period there were several MMs who behaved significantly different from the rest of the group.[21] Nevertheless, rejecting the null hypothesis of equality of the estimates means that the MM has a considerable imbalance between his mandatory sells and buys and his behavior differs from the behavior of other MMs during the particular time period. PIMM tests how different is the average trade imbalance of buys and sells of a particular market maker from the average trade imbalance of buys and sells on the market as a whole. By taking a different position on the balance of buys and sells the market maker is possibly hiding some relevant information from the market or is executing a large block order for its customer.

The second columns of *Tables 7a* and *7b* show the identified time period for the particular stock. To demonstrate the practical use of the method all identification and estimation was done using a 90-day trading window. It is striking that most of the identified periods coincide with significant events or news related to the particular stock. First we discuss the results for Ceske Radiokomunikace (CRA), which was removed from the market in September 2004, although the decision on removal had to be made in 2003. Therefore, our results that MM4 behaved significantly differently from other MMs in the second half of 2003 may suggest that he cooperated with some large informed customer who had better information about the buyout of CRA.

Similarly, the results for Telefonica O_2 resonate with its privatization, indicating that some investors may have been aware of the privatization results and traded on this information ahead of time (see *Figure 2* for a graphical summary of the test: MM7, afternoon). A further example of coincidence with important news is the result for CEZ. As we already mentioned, during 2008 the Czech government was selling nearly 7% of the shares and our results confirm that part of the trade was done using

[20] If the quote is the best available on the market and if some investor reacts to it, the market maker is obliged to execute the trade.

[21] We should point out that the difference does not imply that the market maker is an insider, as he may be processing a large trade order or using dual trading, which is not illegal in the Czech Republic.

Table 7a Extent of Information-Driven Trading Originating from the Behavior of Informed Market Makers

Stock	Time period	morning/ /afternoon	PIN	PIMM	Diff	*T*-stat	*P*-value
CME (MM1)	21.2.2009-4.9.2009	aft	**0.263** (0.029)	**0.352** (0.029)	**0.088** (0.036)	2.46	0.014
CEZ (MM7)	1.4.2009-1.10.2009	morn	**0.253** (0.028)	**0.336** (0.031)	**0.083** (0.036)	2.28	0.022
CEZ (MM4)	1.2.2008-1.7.2008	aft	**0.176** (0.040)	**0.368** (0.029)	**0.193** (0.052)	3.69	0.000
CEZ (MM7)	1.9.2009-9.2.2010	aft	**0.239** (0.037)	**0.338** (0.036)	**0.100** (0.041)	2.43	0.015
CRA (MM4)	26.6.2003-15.10.2003	aft	**0.550** (0.109)	**0.784** (0.088)	**0.234** (0.117)	2.00	0.045
EB (MM1)	25.5.2004-1.11.2004	morn	**0.344** (0.035)	**0.248** (0.038)	**0.097** (0.044)	2.22	0.027
EB (MM1)	1.11.2008-16.5.2009	morn	**0.268** (0.027)	**0.301** (0.028)	**0.033** (0.014)	2.35	0.019
EB (MM7)	1.3.2008-9.9.2008	morn	**0.300** (0.024)	**0.346** (0.025)	**0.046** (0.022)	2.06	0.039
KB (MM7)	5.2.2003-7.7.2003	aft	**0.570** (0.036)	**0.642** (0.034)	**0.072** (0.022)	3.30	0.001
KB (MM7)	2.9.2005-26.1.2006	aft	**0.362** (0.037)	**0.461** (0.036)	**0.100** (0.040)	2.48	0.013
NWR (MM1)	1.7.2008-18.11.2008	aft	**0.301** (0.028)	**0.371** (0.030)	**0.069** (0.018)	3.79	0.000
ORCO (MM5)	5.8.2009-3.12.2009	morn	**0.399** (0.049)	**0.507** (0.043)	**0.107** (0.031)	3.43	0.001
ORCO (MM7)	21.2.2010-21.6.2010	morn	**0.407** (0.058)	**0.584** (0.057)	**0.177** (0.061)	2.90	0.004
ORCO (MM1)	3.3.2009-5.8.2009	aft	**0.466** (0.047)	**0.598** (0.037)	**0.132** (0.041)	3.22	0.001
ORCO (MM7)	1.11.2009-22.4.2010	aft	**0.510** (0.049)	**0.586** (0.051)	**0.076** (0.037)	2.07	0.038

Note: PIMM is the estimate of information-driven trading using the sum of buys and sells excluding the buys and sells of a given market maker. Standard errors are in parentheses.

Source: authors' computations

MM4's market-making activities. Also, the other results for CEZ are connected with important news – during summer 2009 the Czech parliament approved a law on the distribution of free carbon dioxide permits for firms including CEZ.

Another example of important news that coincides with our results is the information about NWR buying a 25% share in Ferrexpo late in 2008. In the case of CME, the uncertain outlook of the firm after the financial crisis and mainly the information about Warner Brothers buying a 31% share in CME is also sufficiently important news to have possibly large orders traded through MM1 during the spring and

Table 7b Extent of Information-Driven Trading Originating from the Behavior of Informed Market Makers

Stock	Time period	morning/ afternoon	PIN	PIMM	Diff	T-stat	P-value
PM (MM7)	1.9.2009-11.2.2010	morn	**0.611** (0.059)	**0.754** (0.046)	**0.143** (0.056)	2.55	0.011
PM (MM7)	21.7.2004-29.11.2004	aft	**0.459** (0.048)	**0.562** (0.047)	**0.103** (0.043)	2.40	0.016
PM (MM7)	11.9.2007-22.2.2008	aft	**0.415** (0.058)	**0.541** (0.056)	**0.127** (0.063)	2.01	0.045
PM (MM2)	1.10.2007-22.2.2008	aft	**0.418** (0.057)	**0.531** (0.057)	**0.113** (0.036)	3.10	0.002
O_2 (MM7)	21.5.2004-31.8.2004	morn	**0.485** (0.050)	**0.638** (0.043)	**0.152** (0.044)	3.50	0.000
O_2 (MM7)	18.5.2005-29.9.2005	morn	**0.527** (0.045)	**0.658** (0.036)	**0.131** (0.034)	3.89	0.000
O_2 (MM7)	1.12.2006-10.4.2007	morn	**0.380** (0.036)	**0.459** (0.037)	**0.080** (0.030)	2.63	0.008
O_2 (MM7)	9.10.2009-27.2.2010	morn	**0.434** (0.043)	**0.528** (0.048)	**0.095** (0.034)	2.79	0.005
O_2 (MM7)	11.6.2004-27.12.2004	aft	**0.466** (0.036)	**0.570** (0.034)	**0.103** (0.034)	3.02	0.003
O_2 (MM7)	21.12.2005-16.5.2006	aft	**0.390** (0.033)	**0.472** (0.038)	**0.082** (0.034)	2.38	0.017
O_2 (MM7)	21.1.2009-29.6.2009	aft	**0.438** (0.033)	**0.510** (0.034)	**0.072** (0.035)	2.09	0.037
O_2 (MM7)	1.1.2010-12.5.2010	aft	**0.449** (0.046)	**0.556** (0.048)	**0.107** (0.053)	2.03	0.043
O_2 (MM10)	14.2.2010-21.7.2010	aft	**0.344** (0.047)	**0.485** (0.045)	**0.141** (0.071)	2.00	0.046
AAA (MM7)	17.5.2009-23.11.2009	morn	**0.464** (0.066)	**0.652** (0.071)	**0.188** (0.081)	2.32	0.020
AAA (MM7)	1.11.2008-22.4.2009	aft	**0.572** (0.073)	**0.682** (0.076)	**0.109** (0.053)	2.06	0.039

Note: PIMM is the estimate of information-driven trading using the sum of buys and sells excluding the buys and sells of a given market maker. Standard errors are in parentheses.

Source: authors' computations

summer of 2009. The financial crisis and the collapse of capital markets decreased the traded volume and therefore also increased the effect of every larger order on the price of the assets. This, together with speculation about whether ORCO would survive its financial problems, might stand behind our results of the significantly different behavior of MM5, MM1, and MM7 during 2009–2010 for ORCO.

Overall, the results confirm the perception that a high percentage of block trades (around 30%) or large orders might have a significant impact on the behavior of some MMs, as you can see on the results for CEZ, Komercni banka, and Phillip

Table 8 Extent of Information-Driven Trading Before and After Changing the Lot Size

Stock	Date	LOT 1	LOT 2	morn/ /aft	PIN 1	PIN 2	Diff	*T*-stat	*P*-value
CEZ	15.10.04	20,000	10,000	morn	**0.462**	**0.352**	**0.110**	1.52	0.128
					(0.038)	(0.030)	(0.072)		
				aft	**0.517**	**0.416**	**0.101**	1.55	0.120
					(0.042)	(0.032)	(0.065)		
CEZ	12.8.05	10,000	5,000	morn	**0.333**	**0.232**	**0.101**	2.24	0.025
					(0.027)	(0.030)	(0.045)		
				aft	**0.428**	**0.332**	**0.096**	1.48	0.139
					(0.041)	(0.030)	(0.065)		
EB	19.9.03	500	1,000	morn	**0.376**	**0.349**	**0.027**	0.47	0.636
					(0.043)	(0.037)	(0.057)		
				aft	**0.500**	**0.521**	**-0.021**	0.25	0.802
					(0.061)	(0.048)	(0.084)		
EB	12.3.04	1,000	500	morn	**0.384**	**0.362**	**0.022**	0.33	0.744
					(0.037)	(0.035)	(0.066)		
				aft	**0.538**	**0.434**	**0.104**	1.29	0.197
					(0.047)	(0.043)	(0.080)		
KB	5.9.03	2,000	1,000	morn	**0.523**	**0.352**	**0.171**	3.02	0.003
					(0.030)	(0.032)	(0.057)		
				aft	**0.575**	**0.446**	**0.130**	1.95	0.051
					(0.036)	(0.034)	(0.066)		
KB	30.6.08	1,000	500	morn	**0.310**	**0.281**	**0.029**	0.80	0.426
					(0.031)	(0.027)	(0.036)		
				aft	**0.367**	**0.266**	**0.100**	1.79	0.073
					(0.038)	(0.030)	(0.056)		
PM	12.3.04	200	100	morn	**0.741**	**0.498**	**0.243**	2.68	0.007
					(0.040)	(0.037)	(0.091)		
				aft	**0.713**	**0.489**	**0.225**	3.20	0.001
					(0.047)	(0.047)	(0.070)		
UNI	24.2.05	20,000	10,000	morn	**0.476**	**0.334**	**0.142**	2.23	0.025
					(0.050)	(0.050)	(0.064)		
				aft	**0.380**	**0.278**	**0.102**	1.45	0.148
					(0.060)	(0.059)	(0.070)		
ZEN	29.6.07	3,000	2,000	morn	**0.287**	**0.278**	**0.009**	0.21	0.837
					(0.032)	(0.044)	(0.043)		
				aft	**0.299**	**0.144**	**0.154**	1.93	0.054
					(0.037)	(0.061)	(0.080)		

Note: The table shows the extent of information-driven trading within 90 trading days before and after a change in the lot size. Standard errors are in parentheses.

Source: authors' computations

Morris during 2003–2005. Although the percentage of block trades decreased remarkably in 2006, the behavior of MMs probably did not change, as the SPAD trades with no identification experienced a significant increase at the same time. Our results

suggest that even though market participants might be aware of the different behavior of several MMs, they are not able to compete with them due to superior information coming, for example, from detailed information about large orders.

Finally, we focused on the effect of changes in the trading lot size on trading behavior and on PIN. As we already mentioned, changing the lot size may affect the extent of information-driven trading, as informed traders are more likely to trade larger volumes. Smaller lot volumes might attract more uninformed investors. The estimation and test results are summarized in *Table 8*.

As we can see, most of the changes in the lot size significantly affected the extent of information-driven trading, as lot breakups attracted more retail and therefore uninformed investors. Overall, a smaller lot size means more trades with the particular stock, higher attractiveness for individual investors, and a lower extent of PIN. The exception is Erste bank (EB), whose shares did not significantly react to either an increase or decrease in the lot size. In terms of information-driven trading, this is still consistent, since the primary market of EB is the Vienna Stock Exchange and trading in Prague is much smaller compared to its main market.

6. Conclusion

In this paper we analyze the behavior of MMs and the ability to maintain private information about large orders. We propose an automatic procedure using order flows to detect and test specific positions of particular MMs in an electronic dealers market. Trading data with one side of the mandatory buy/sell trade orders identified are used to demonstrate our method.

We found significant differences in behavior among MMs on the Prague Stock Exchange, supporting the perception that they play a dominant role in affecting the price for a short time interval as well as for a longer period. Although the other participants in the market may be aware that some of the MMs might possess private information about the value of the asset, they are not able to reveal the full information. Further, our analysis confirms that important changes such as decreasing the volume of the trading lot may affect (decrease) the extent of order flow information--driven trading.

From the trading perspective it could be understood that on a thin market MMs should be somehow protected to be able to maintain private information about their sizable (block) orders and be able to face the threats of predatory trading and increased volatility during such trades. Nevertheless, the current practice of MMs in the Prague Stock Exchange could threaten minority and uninformed investors because prices can then no longer convey all relevant information. This observation leads to the conclusion that further regulation might be beneficial. However, the optimal policy from the regulatory point of view is not so straightforward and is beyond the scope of the present paper. From our estimation it is clear, though, that increased regulation by introducing so-called trading books (detailed recording of all trades conducted) for each broker does not help. As we mentioned (forbidden) block trades were basically transformed into SPAD trades with missing party identification. This might suit some of the large institutional players, but we believe that it results in a less transparent market with lower trading volume and lower informational content. All of these results could be harmful, especially on a market with a significant pres-

ence of foreign investors. This raises the question about introducing an upstairs market (similar to the NYSE, for example), a transparent trading segment that would serve institutional investors, and keeping a variant of SPAD with smaller lot size suitable for small individual investors. It is clear from our results that an appropriate (i.e., much smaller) trading lot size would reduce the amount of information-driven trading and would also increase the trading volume and attract small individual investors.

In general, given that this study uses an automatic procedure, has only modest assumptions, and includes a model that is relatively easy to use, we believe that the methodology in this paper could be used by investors as well as regulatory authorities on emerging markets to identify unusual order flows and/or suspicious behavior by particular market participants.

REFERENCES

Anand A, Chakravarty S, Martell T (2005): Empirical evidence on the evolution of liquidity: Choice of market versus limit orders by informed and uninformed traders. *Journal of Financial Markets*, 8(3):289–309.

Barclay MJ, Warner JB (1993): Stealth trading and volatility: Which trade move prices? J*ournal of Financial Economics*, 34(3):281–205.

Barclay MJ, Hendershott T, McCormick DT (2003): Competition among trading venues: Information and trading on electronic communications networks. *Journal of Finance*, 8(6):2637–2665.

Boehmer E (2005): Dimensions of execution quality: Recent evidence for US equity markets. *Journal of Financial Economics*, 78(3):553–582.

Boehmer E, Grammig J, Theissen E (2007): Estimating the probability of informed trading – does trade misclassification matter? *Journal of Financial Markets*, 10(1):26–47.

Brunnermeier MK, Pederssen LH (2005): Predatory trading. *Journal of Finance*, 60(4):1825–1863.

Chakravarty S (2001): Stealth-trading: Which traders' trades move stock prices? *Journal of Financial Economics*, 61(2):289–307.

Chakravarty S, Li K (2003): A Bayesian analysis of dual trader informativeness in futures markets. *Journal of Empirical Finance*, 10(3):355–371.

Chan LKC, Lakonishok J (1995): The Behavior of Stock Prices Around Institutional Trades. *Journal of Finance*, 50(4):1147–1174.

Charoenwong C, Ding DK, Jenwittayaroje N (2010): Price movers on the Stock Exchange of Thailand: Evidence from a fully automated order-driven market. *Financial Review*, 45:761–783.

Easley D, Hvidkjaer S, O'Hara M (2010): Factoring Information into Returns. *Journal of Financial and Quantitative Analysis*, 45(2):293–309.

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

Fishman MJ, Longstaff F (1992): Dual trading in futures markets. *Journal of Finance*, 47(2): 643–671.

Garfinkel JA, Nimalendran M (2003): Market Structure and Trader Anonymity: An analysis of insider trading. *Journal of Financial and Quantitative Analysis,* 38(3):591–610.

Golec J (2007): Are the Insider Trades of a Large Institutional Investor Informed? *Financial Review,* 42(2):161–190.

Grammig J, Schiereck D, Theissen E (2001): Knowing me, knowing you: Trader anonymity and informed trading in parallel markets. J*ournal of Financial Markets*, 4(4):385–412.

Hanousek J, Podpiera R (2002): Information Driven Trading at The Prague Stock Exchange: Evidence from Intra-day Data. *Economics of Transition*, 10(3):747–759.

Hanousek J, Podpiera R (2004): Czech Experience with Market Maker Trading System. *Economic Systems*, 28(1):177–91.

Heidl HG, Huang RD (2002): Information-based trading in dealer and auction markets: An analysis of exchange listings. *Journal of Financial and Quantitative Analysis*, 37(3):391–424.

Keim DB, Madhavan A (1995): Anatomy of the trading process Empirical evidence on the behavior of institutional traders. *Journal of Financial Economics*, 37(3):371–398.

Kurov A, Lasser D (2004): Price dynamics in the regular and E-mini futures markets, *Journal of Financial and Quantitative Analysis*, 39(2):365–384.

Kyle A (1985): Continuous auctions and insider trading. *Econometrica*, 53(6):331–359.

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

Lee J, Yi CH (2001): Trade size and information-motivated trading in the options and stock markets. *Journal of Financial and Quantitative Analysis*, 36(4):485–501.

Madhavan A, Panchapagesan V (2000): Price discovery in auction markets: a look insided the black box. *Review of Financial Studies*, 13(3):627–658.

Roell A (1990): Dual capacity trading and the quality of the market. *Journal of Financial Intermediation*, 1(2):105–124.

Rogers WH (1993): sg16.4: Comparison of nbreg and glm for negative binomial. *Stata Technical Bulletin*, 16(7). Reprinted in: *Stata Technical Bulletin Reprints*, 3:82–84. College Station, TX: Stata Press.

Sarkar A (1995): Dual trading: winners, losers and market impact. *Journal of Financial Intermediation*, 4(1):77–93.

Sherwood A (1997): The impact of electronic trading on regulatory surveillance. *Journal of Financial Regulation and Compliance*, 5(4):299–305.

Vega C (2006): Stock price reaction to public and private information. *Journal of Financial Economics*, 82(1):103–133.

Wang A, Chae J (2003): Who makes markets? Cross-sectional variation in voluntary liquidity provision. MIT Sloan School of Management, *MIT, Cambridge, Working paper*, no. 4434-03.