



Informed trading, order flow shocks and the cross section of expected returns in Borsa Istanbul

Murat Tiniç & Aslihan Salih

To cite this article: Murat Tiniç & Aslihan Salih (2019): Informed trading, order flow shocks and the cross section of expected returns in Borsa Istanbul, Applied Economics, DOI: [10.1080/00036846.2019.1676386](https://doi.org/10.1080/00036846.2019.1676386)

To link to this article: <https://doi.org/10.1080/00036846.2019.1676386>



Published online: 09 Oct 2019.



Submit your article to this journal [↗](#)



Article views: 40



View related articles [↗](#)



View Crossmark data [↗](#)



Informed trading, order flow shocks and the cross section of expected returns in Borsa Istanbul

Murat Tiniç^{a,b} and Aslihan Salih^c

^aDepartment of Management, Bilkent University, Ankara, Turkey; ^bFaculty of Management, Kadir Has University, Istanbul, Turkey; ^cFaculty of Economics and Administrative Sciences, TED University, Ankara, Turkey

ABSTRACT

This paper examines the relationship between information asymmetry and stock returns in Borsa Istanbul. For all stocks that are traded in Borsa Istanbul between March 2005 and April 2017, we estimate the probability of informed trading (PIN) to proxy for information asymmetry. Firm-level cross-sectional regressions indicate a statistically insignificant relationship between PIN estimates and future returns. Moreover, univariate and multivariate portfolio analyses assert that investors that hold stocks that have high information asymmetry do not obtain significant future returns. Consequently, our results suggest that information asymmetry proxied by PIN is a firm-specific risk and can be eliminated with portfolio diversification. Findings are robust to different factorizations in estimating PIN and free of any bias due to trade classification algorithms, boundary solutions, floating-point exceptions and symmetric order flow shocks.

KEYWORDS

Information asymmetry; market microstructure; borsa istanbul; asset pricing

JEL CLASSIFICATION

G11; G12

I. Introduction

There are conflicting views on the relationship between information risk which is arising from the information asymmetry between informed and uninformed investors, and the securities returns. Easley and O'hara (2004) construct a rational expectations model with information asymmetry to show that uninformed investors are inferior in adjusting their portfolios quickly, in the presence of new information. Therefore, *ceteris paribus*, uninformed investors demand a premium for holding stocks that have a high presence of informed traders which in turn implies that information risk is systematic and undiversifiable. On the contrary, Hughes, Liu, and Liu (2007) argue that in a large economy where the number of securities is infinite, increase in the information asymmetry yields an increase in returns for the overall market. Their claims indicate that after controlling for systematic factors, firm-specific information should not affect the cross-section of expected returns and thus the information risk is idiosyncratic and fully diversifiable.

These conflicting views have empirical implications. Easley et al. (1996) is a cornerstone study that quantifies the information asymmetry for a stock trade. Easley et al. (1996) and Easley, Hvidkjaer,

and O'Hara (2002) derive a proxy called the probability of informed trading (PIN) which aims to measure the proportion of informed investors. Easley et al. (1996) show that transaction costs are, on average, larger for stocks that have higher levels of the PIN. Similarly, Easley, Hvidkjaer, and O'Hara (2002) and Easley, Hvidkjaer, and O'Hara (2010) indicate that investors demand a premium for holding stocks that have higher PIN levels.

On the contrary, Lai, Ng, and Zhang (2014) show that the information is not a priced risk for 47 different countries. More importantly, Duarte and Young (2009) decompose PIN into two separate components, one related to adverse selection (adjusted PIN), the other is related to stocks liquidity. They adjust the PIN measure by introducing the probability of systematic order-flow shocks (PSOS) that jointly increase the arrival of buy and sell order flows. Their results also indicate that PSOS is a priced risk rather than the adjusted PIN.

There is a growing literature on the role of information asymmetry on the stocks traded in the Turkish stock market, Borsa Istanbul (BIST). The microstructure analysis of Tiniç and Savaser (2018) examine the differences in information between foreign and domestic investors. They

show that local investors have a price impact over a broader array of stocks however the foreign price impact significantly increases after June 2013 with the increased political turmoil within the country. They also indicate that foreign impact clusters over stocks with larger market capitalization. Moreover, Simsir and Simsek (2018) show an abnormal trading volume around the first private notification times which is right before public disclosure announcements made public in BIST. They argue that the profits from the trades right before an announcement made public can sum up to 77 million \$. To this end, this paper is an attempt to examine the systematic relationship between information asymmetry and stock return in BIST. Due to the lack of studies that examine the stock mispricing and the microstructure of BIST, this paper fills an essential gap in the literature.

In this paper, we analyse the nature of expected returns of high-PIN stocks traded in Borsa Istanbul to explore the preferences of investors towards the risk of trading with informed investors. We first examine this relationship through Fama–MacBeth regressions (Fama and MacBeth 1973). Besides, similar to Easley, Hvidkjaer, and O’Hara (2010), we examine the behaviour of PIN-sorted portfolios through cross-sectional regressions.

Our results are in line with the theoretical expectations of Hughes, Liu, and Liu (2007) and the empirical findings of Lai, Ng, and Zhang (2014). We fail to reject the hypothesis that the information asymmetry is an idiosyncratic risk in BIST. First, we do not observe any predictive power of PIN on explaining the next month returns in firm-level cross-sectional regressions. Following this, our univariate and multivariate portfolio analyses assert that the investors who trade on BIST do not expect a premium for holding stocks that have high levels of information asymmetry. The results are free from any bias due to classification algorithms, floating-point exception, boundary solutions and systematic order flow shocks in estimating PIN.

The structure of the paper is as follows; in the next section, we provide a brief literature review on the PIN model and its extensions, Section III introduces the sample formation, Section IV discusses the methodology, Section V presents the results, Section VI concludes.

II. PIN model and its extensions

The sequential trading model of Easley et al. (1996) assumes that for a stock on a trading day t , an information event occurs with probability α . This information event can be a bad news with probability δ and it can be a good news with probability $(1 - \delta)$. Informed investors are assumed to interpret the content of the news perfectly. That is, they are assumed to arrive at market to buy only when there is good news and they arrive at market to sell only when there is bad news. Informed investors assumed to follow a Poisson distribution with parameter μ . Uninformed investors arrive at the market independent of the information event and the content. They are assumed to arrive at market to place a buy (sell) order with Poisson distribution with parameter ϵ_b (ϵ_s).

Then, for given trading day t , the joint probability distribution for observing buy and sell orders (B_t, S_t) given the parameter vector $\Theta = \{\alpha, \delta, \mu, \epsilon_b, \epsilon_s\}$ can be written as follows:

$$\begin{aligned} f(B_t, S_t | \Theta) \equiv & \alpha \delta \exp(-\epsilon_b) \frac{\epsilon_b^{B_t}}{B_t!} \exp[-(\epsilon_s + \mu)] \frac{(\epsilon_s + \mu)^{S_t}}{S_t!} \\ & + \alpha(1 - \delta) \exp[-(\epsilon_b + \mu)] \frac{(\epsilon_b + \mu)^{B_t}}{B_t!} \exp(-\epsilon_s) \frac{\epsilon_s^{S_t}}{S_t!} \\ & + (1 - \alpha) \exp(-\epsilon_b) \frac{\epsilon_b^{B_t}}{B_t!} \exp(-\epsilon_s) \frac{\epsilon_s^{S_t}}{S_t!} \end{aligned} \quad (1)$$

The parameter estimates $\hat{\Theta} = \{\hat{\alpha}, \hat{\delta}, \hat{\mu}, \hat{\epsilon}_b, \hat{\epsilon}_s\}$ can be obtained by maximizing the joint log-likelihood function given the order input matrix (B_t, S_t) over T trading days. The non-linear objective function of this problem can be written as;

$$L(\Theta | T) \equiv \sum_{t=1}^T L(\Theta | (B_t, S_t)) = \sum_{t=1}^T \log[f(B_t, S_t | \Theta)] \quad (2)$$

The maximization problem is subject to boundary constraints $\alpha, \delta \in [0, 1]$ and $\mu, \epsilon_b, \epsilon_s \in [0, \infty)$. The PIN estimate is then the proportion of informed investors in the market which is given by;

$$\widehat{PIN} = \frac{\hat{\alpha} \hat{\mu}}{\hat{\alpha} \hat{\mu} + \hat{\epsilon}_b + \hat{\epsilon}_s} \quad (3)$$

Floating-point exception and different likelihood factorizations

The likelihood factorization presented in Equation (1) (EKOP factorization) is prone to selection bias. Lin and Ke (2011) show that the feasible solution set for the non-linear optimization problem shrinks significantly as the daily number of buy and sell orders increases. The optimal value for the log-likelihood in Equation (2) becomes undefined, for large enough (B_t, S_t) whose factorials cannot be computed by mainstream computers (floating-point exception (FPE)). Therefore, studies that use EKOP factorization may systematically exclude stocks that are actively traded.

To deal with this problem Duarte and Young (2009) modifies the EKOP factorization by replacing $\exp(-\epsilon_b) \frac{\epsilon_b^{B_t}}{B_t!}$ with $\exp[-\epsilon_b + B_t \ln(\epsilon_b) - \sum_{i=1}^{B_t} \ln(i)]$ and other components of the likelihood with the similar modifications. In addition, two other factorizations that are algebraically equivalent to (1) are provided in order to avoid FPE. The first factorization (EHO factorization) is provided by Easley, Hvidkjaer, and O'Hara (2010). The other (LK factorization) is provided by Lin and Ke (2011). Both studies modify the factorization such that $B_t!$ and $S_t!$ become constant with respect to the parameter vector Θ and dropped in the maximum likelihood procedure.

Boundary solutions and different estimation methods

Yan and Zhang (2012) show that the parameter estimates $\hat{\alpha}$ and $\hat{\delta}$ usually fall onto the boundaries of the parameter space $[0,1]$ when estimating PIN. Yan and Zhang (2012) show that this is due to the fact that the MLE algorithm gets stuck at a local optimum. Notice that the PIN estimates presented in Equation (3) are directly related to the estimate of $\hat{\alpha}$. Letting $\hat{\alpha}$ equal to zero will directly equate the PIN estimate to zero as well. This can be problematic on studies that estimate PIN over a certain period in which it is certain that an information event has occurred. For example, for studies that estimate PIN over a quarter, one can be sure that there is at least one information event, earnings announcement.

To overcome this problem, Yan and Zhang (2012) propose a grid-search algorithm (YZ algorithm) that spans the parameter space for 125

different sets of initial parameter values. For different initial parameter values, YZ algorithm obtains different PIN estimates and selects the estimate that gives the highest likelihood value for non-boundary solutions.

However, YZ algorithm is quite time-consuming. In recent years, clustering algorithms become increasingly popular in estimating PIN, due to efficiency concerns. Gan, Wei, and Johnstone (2015) and Ersan and Alici (2016) use clustering algorithms to estimate the PIN. They use the mean difference in the order imbalance to group the data into three groups, no news, good news, bad news. Once each observation is grouped into their no news, good news or bad news clusters, the parameter estimates are obtained simply by counting. Celik and Tiniç (2018) provide a detailed survey on the specifications and performance of the methods presented above.

Probability of symmetric order-flow shock (PSOS)

The sequential trading model of Easley et al. (1996) enforces a negative contemporaneous correlation between the number of buy and sell orders (B, S) . Informed trading can inflate the number of buy and sell orders, separately. Duarte and Young (2009) indicate that the contemporaneous covariance between the number of buy and sell orders in PIN model is given as follows:

$$\text{cov}(B, S) = (\alpha\mu)^2(\delta - 1)\delta \leq 0 \quad (4)$$

However, Duarte and Young (2009) argue that this aspect of the model is not observed empirically. They show a strong positive correlation between buy and sell orders and high levels of buy and sell volatility.

To account for the different levels of variance between buy and sell orders, they separate the arrival process for informed traders to place buy and sell orders from one another. They defined μ_b (μ_s) to be the arrival rate of informed traders to place a buy (sell) order. More importantly, they allow for increased buy and sell variation along with a positive correlation between buys and sells as they define another event that causes both buy and sell order flow to increase. They called this event a *symmetric order-flow shock*. The probability of such an event is assumed to follow a Bernoulli distribution with parameter γ conditional on the

absence of private information. The probability of such an event conditional on the arrival of private information is also assumed to follow a Bernoulli distribution, but with a different parameter, γ' . In case of a symmetric order flow shock, the additional arrival rate of buys (sells) is defined as ϕ_b (ϕ_s).

Then, the parameter vector of the extended model becomes, $\Theta = \{\alpha, \delta, \gamma, \gamma', \mu_b, \mu_s, \epsilon_b, \epsilon_s, \phi_b, \phi_s\}$.

The joint probability function in Equation (1) can be updated as follows:

$$\begin{aligned}
f(B_t, S_t | \Theta) &\equiv p(B, S) \\
&\equiv (1 - \alpha)(1 - \gamma) \exp[-\epsilon_b + B_t \ln(\epsilon_b)] \\
&\quad - \sum_{i=1}^{B_t} \ln(i) \exp[-\epsilon_s + S_t \ln(\epsilon_s)] \\
&\quad - \sum_{i=1}^{S_t} \ln(i) + (1 - \alpha)\gamma \exp[-(\epsilon_b + \phi_b)] \\
&\quad + B_t \ln(\epsilon_b + \phi_b) - \sum_{i=1}^{B_t} \ln(i) \exp[-(\epsilon_s + \phi_s)] \\
&\quad + S_t \ln(\epsilon_s + \phi_s) - \sum_{i=1}^{S_t} \ln(i) \\
&\quad + \alpha(1 - \delta)(1 - \gamma') \exp[-(\epsilon_b + \mu_b)] \\
&\quad + B_t \ln(\epsilon_b + \mu_b) - \sum_{i=1}^{B_t} \ln(i) \exp[-\epsilon_s + S_t \ln(\epsilon_s)] \\
&\quad - \sum_{i=1}^{S_t} \ln(i) + \alpha(1 - \delta)\gamma' \exp[-(\epsilon_b + \mu_b + \phi_b)] \quad (5) \\
&\quad + B_t \ln(\epsilon_b + \mu_b + \phi_b) - \sum_{i=1}^{B_t} \ln(i) \exp[-(\epsilon_s + \phi_s)] \\
&\quad + S_t \ln(\epsilon_s + \phi_s) \sum_{i=1}^{S_t} \ln(i) \\
&\quad + \alpha\delta(1 - \gamma') \exp[-\epsilon_b + B_t \ln(\epsilon_b)] \\
&\quad - \sum_{i=1}^{B_t} \ln(i) \exp[-(\epsilon_s + \mu_s) + S_t \ln(\epsilon_s + \mu_s)] \\
&\quad - \sum_{i=1}^{S_t} \ln(i) + \alpha\delta\gamma \exp[-(\epsilon_b + \phi_b) + B_t \ln(\epsilon_b + \phi_b)] \\
&\quad - \sum_{i=1}^{B_t} \ln(i) \exp[-(\epsilon_s + \phi_s + \mu_s)] \\
&\quad + S_t \ln(\epsilon_s + \phi_s + \mu_s) - \sum_{i=1}^{S_t} \ln(i)
\end{aligned}$$

The structural model presented in Equation (5) enables a positive covariance between buy and sell order. The probability of informed trading adjusted for the symmetric order flow shock is given by;

$$\begin{aligned}
AdjPIN &\equiv \frac{\alpha(\delta\mu_s + (1 - \delta)\mu_b)}{\alpha(\delta\mu_s + (1 - \delta)\mu_b) + (\phi_b + \phi_s)(\alpha\gamma' + (1 - \alpha)\gamma) + \epsilon_b + \epsilon_s} \quad (6)
\end{aligned}$$

The probability of symmetric order flow shock (PSOS) is then given by;

$$\begin{aligned}
PSOS &\equiv \frac{(\phi_b + \phi_s)(\alpha\gamma' + (1 - \alpha)\gamma)}{\alpha(\delta\mu_s + (1 - \delta)\mu_b) + (\phi_b + \phi_s)(\alpha\gamma' + (1 - \alpha)\gamma) + \epsilon_b + \epsilon_s} \quad (7)
\end{aligned}$$

The problem of boundary solutions are also evident for estimating PIN when DY factorization is used. To this end, we choose 10 different random starting points for $(\alpha^0, \delta^0, \gamma^0, \gamma'^0)$ from uniform distribution $U(0, 1)$ for parameters that are bounded from above and below. For parameters that are bounded only from above, $(\mu_A^0, \mu_S^0, \epsilon_A^0, \epsilon_S^0, \phi_A^0, \phi_S^0)$ we set the arrival rates to be the one-third of the respective buy and sell values within the specified period.

III. Sample formation

We collect the limit order book for all stocks traded in BIST National Markets Listing, between March 2005 to April 2017. Our sample consists of 498 different stocks. We obtain price and accounting information from Bloomberg Terminal. We calculate the firm-specific variables such as size (SIZE), book-to-market ratio (BTM), beta (BETA), illiquidity (ILLIQ), idiosyncratic volatility (IVOL), reversal (REV), maximum daily return (MAX) and momentum (MOM) at monthly frequency as shown in the Appendix. Each order in the limit order book contains the date, time, stock ticker, order ID, order type, quantity and price stamps. Order type enables us to track the order process for buys and sells, given a trading day. Therefore, we construct the order input matrix (B_t, S_t) from the actual number of buy and sell orders, as a panel data which has trading days in the time dimension and each firm in BIST in the cross-section. We exclude all orders that arrive in the opening auction.

The observation of the actual order arrivals frees us from considering the trade initiation sequence. This freedom relieves us from choosing different trade

classification algorithms to assign values for (B_t, S_t) . The literature on the accuracy of algorithms that classify trades as buyer or seller initiated is inconclusive. However, all classification algorithms are shown to create some bias in BIST (Aktas and Kryzanowski 2014). Our results presented in the next section, are therefore robust to any biases that might arise due to improper selection of trade classification algorithms.

Table 1 presents the descriptive statistics on the daily number of buy and sell orders. Daily buy orders are more volatile compared to sell orders. More importantly, the positive correlation between buy and sell orders is evident for BIST. We show in Table 1 that the minimum cross-sectional correlation between buy and sell orders is 10% throughout our sample period. The average correlation is 85%. Similar, Duarte and Young (2009), we do not observe any negative correlation among buy and sell value. This implies that the Duarte and Young (2009) extension of the PIN model which takes the probability of symmetric order flow shock into consideration is necessary when examining the pricing implications of information asymmetry in BIST.

Table 2 shows the descriptive statistics for the PIN and PSOS estimates calculated using the Duarte and Young (2009) factorization along with the descriptive statistics for firm-specific variables. Table 3 gives the pair-wise correlations among firm-specific variables. In line with Alkan and Guner (2018), the highest correlation is between IVOL and MAX factors with 84%. This is followed by the correlation between RETURN and MAX with 34%. The largest negative correlation is among PIN and PSOS factors.

IV. Methodology

Fama–Macbeth regressions

To test the information asymmetry hypothesis, we estimate the cross-sectional regression presented in (8) for each month in our sample. For each month in our sample the returns are regressed on PIN and various firm-specific measures.

$$R_{i,t+n} = \gamma_{0,t} + \gamma_{1,t}PIN_{i,t} + \gamma_{2,t}BETA_{i,t} + \gamma_{3,t}SIZE_{i,t} + \gamma_{4,t}BTM_{i,t} + \gamma_{5,t}X_{i,t} + \epsilon_{i,t} \quad (8)$$

where $R_{i,t}$ is the return on stock i , $i \in \{1 \dots 498\}$ for holding periods of n months after month t . We try three different holding periods, 1-month, 3-months, 6-months, in line with Atilgan, Demirtas, and Gunaydin (2016); Alkan and Guner (2018). The vector $X_{i,t}$ includes other firm-specific factors such as MOM, MAX, IVOL, PSOS and ILLIQ which are described in the Appendix. We form our first hypothesis as follows:

Hypothesis 1

$$H_0 : \gamma_{1,t} = 0$$

$$H_A : \gamma_{1,t} > 0$$

Univariate portfolio analysis

Information asymmetry hypothesis of Easley and O'hara (2004) suggests that uninformed investors demand compensation for holding stocks that have high information asymmetry as they cannot update their portfolios to the new information arrive at the

Table 1. Descriptive statistics on number of buy and sell orders.

Statistics	Mean Buy	Mean Sell	Variance Buy	Variance Sell	Correlation
Minimum	18.56	14.11	17.71	13.53	0.10
25th Percentile	248.97	182.06	355.83	229.35	0.81
Median	411.84	290.13	562.06	343.73	0.88
Mean	521.07	398.34	733.13	487.61	0.85
75th Percentile	620.56	445.17	856.33	557.30	0.92
Maximum	4627.44	4223.77	5728.53	2945.22	0.98

In this table, we present the descriptive statistics for the cross-sectional distribution of a series of statistics based on the daily number of buy and sell observations, (B_t, S_t) for each stock in our sample. The buy and sell orders are observed in our data set which covers all stocks that are traded on BIST between March 2005 and April 2017. Our sample consists of 498 different stocks. For each stock, in each trading day, we compute the total number of buy and sell orders submitted to BIST from the intraday limit order book. Then, for each stock, we compute the mean and variance of the daily buy and sell orders throughout our sample period. First two columns represent the descriptive statistics for the mean buy and sell values in our cross-section. Similarly, third and fourth column represents the descriptive statistics for the variance of the daily buy and sell values. Final column represents the descriptive statistics for the Spearman correlation coefficient between daily number of buy and sell orders.

Table 2. Descriptive statistics on the cross-sectional distributions of monthly factors.

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
RETURN	44967	0.01	0.14	-0.89	2.94
BETA	44957	0.60	0.56	-13.07	23.33
SIZE	44967	5.25	1.95	0.29	10.68
BTM	44967	1.10	1.23	0.00	53.40
MOM	44226	-0.05	0.62	-73.01	0.95
REV	44501	0.01	0.14	-0.89	2.94
ILLIQ	44967	0.03	1.10	0.00	100.15
MAX	44967	0.06	0.04	-0.10	0.20
IVOL	44957	0.02	0.01	0.00	0.18
PIN	44966	0.16	0.17	0.00	1.00
PSOS	44966	0.24	0.25	0.00	1.00

In this table, we present the descriptive statistics for firm-specific factors that are calculated for all stocks that are traded in BIST between March 2005 – April 2017. RETURN represents the monthly percentage returns. BETA is the systematic risk factor. SIZE is the logarithm of the end of month market capitalization. BTM is the book-to-market ratio. MOM is the momentum variable. REV represents the reversal variable. ILLIQ is the illiquidity measure. MAX is the maximum daily return within a month. IVOL is the idiosyncratic volatility. PIN is the probability of informed trading estimated using DY factorization. PSOS is the probability of systematic order flow shocks.

Table 3. Factor correlations.

	RETURN	BETA	SIZE	BTM	MOM	REV	ILLIQ	MAX	IVOL	PIN	PSOS
RETURN	1										
BETA	-0.109	1									
SIZE	0.058	0.174	1								
BTM	-0.080	-0.011	-0.233	1							
MOM	-0.010	0.006	0.132	-0.210	1						
REV	0.032	0.007	0.035	-0.036	0.026	1					
ILLIQ	-0.006	0.009	0.070	0.001	-0.001	-0.002	1				
MAX	0.343	0.105	-0.164	0.008	-0.043	0.012	0.000	1			
IVOL	0.202	-0.013	-0.254	0.050	-0.050	-0.002	0.049	0.841	1		
PIN	0.027	0.017	0.034	-0.014	0.011	0.009	-0.004	0.029	0.024	1	
PSOS	-0.030	0.066	0.130	0.002	0.000	0.010	0.003	-0.054	-0.047	-0.217	1

This table presents the cross-sectional correlations among time-series averages of factors. PIN is the probability of informed trading. PSOS is the probability of systematic order flow shocks. Size is the market capitalization. BETA is the market risk factor. MOM is cumulative return over the period $[t-7, t-1]$. MAX is the maximum daily return within the month t . IVOL is the idiosyncratic volatility. BTM is the book to market ratio. ILLIQ is the Amihud (2002) illiquidity measure.

market. To examine the relationship between information asymmetry and stock returns, we use portfolio analysis where we form deciles each month by sorting stocks based on their PIN estimate. In this setting, the stocks that are in the lowest decile mimics an investor's portfolio which is predominantly composed of stocks that have low information asymmetry. Contrarily, the stocks that are in the highest decile mimics an investors portfolio which is composed solely of stocks that have high information asymmetry. We then mimic an investors portfolio who has a long position in stocks with high information asymmetry and short position of equal size in stocks that have low information asymmetry and check whether the resulting zero investment portfolio gains significant returns. Following Easley, Hvidkjaer, and O'Hara (2002), Easley and O'hara (2004), and Easley, Hvidkjaer, and O'Hara (2010) we expect that zero investment portfolio will earn a significant a positive return. Letting R^{High} and

R^{Low} denote the returns of high-PIN and low-PIN stocks, our second hypothesis is as follows:

Hypothesis 2:

$$H_0 : R^{High} - R^{Low} = 0$$

$$H_A : R^{High} - R^{Low} > 0$$

Multivariate portfolio analysis

Easley, Hvidkjaer, and O'Hara (2010) documents that informed trading is larger for small firms. In Table 3, we observe a negative correlation between PIN estimates and SIZE. Although this finding is contradicting with Easley, Hvidkjaer, and O'Hara (2010), it is in line with Tiniç and Savaser (2018) shows that informed trading is larger for firms that are larger. This relationship might arise from the fact that larger firms, due to their nature, produce more information through investment or research and development activity.

To this end, in addition to univariate portfolio analysis, we apply multivariate portfolio analysis in which we consider the variation in firm size. Within the scope of multivariate portfolio analysis, first, we sort the stocks into deciles based on their market value at the end of previous month. Then, we divide each size decile into three groups based on the PIN estimates of the stocks that are in that decile. This procedure gives us 30 different portfolios. We then follow the returns of these portfolios for one, three and 6 months. We form zero investment portfolios R_p for each decile ($p = 1 \dots 10$) by subtracting the returns of the high-PIN portfolio (R_p^{High}) from the low-PIN (R_p^{Low}) portfolio. These returns provides us to examine the impact of information asymmetry controlling for firm value.

Similarly, controlling for firm value, we test the relationship between information asymmetry and stock returns as follows:

Hypothesis 3:

$$H_0 : R_p = 0$$

$$H_A : R_p > 0$$

Factor tests

Finally, we use factor tests to examine the relationship between information asymmetry and stock returns under the control of size and profitability factors that Fama and French (1993) introduced for U.S. markets. Akdeniz, Altay-Salih, and Aydoğan (2000) show that size and profitability are systematic risk factors for BIST. The standard model is presented as below;

$$R_p = \vartheta_p^1 + \vartheta_p^2 R_m + \vartheta_p^3 SMB + \vartheta_p^4 HML + \epsilon_p \quad (9)$$

$$p = 1 \dots 10$$

where R_m represents the market (BIST 100 Index) return series, SMB and HML are the size and value factors of Fama and French (1993) applied for all stocks traded in BIST. We update SMB and HML portfolios yearly, in order to be consistent with the

literature. We then test the hypothesis that whether *the constant terms* $\vartheta^1 \equiv (\vartheta_1^1 \dots \vartheta_{10}^1)$ is jointly equal to zero or not.

Hypothesis 4:

$$H_0 : \vartheta^1 = 0$$

$$H_A : \vartheta^1 > 0$$

To test this hypothesis, we use the F statistics of (Gibbons, Ross, and Shanken 1989) (GRS test statistic).

V. Results

Regression analysis

We present Fama–MacBeth regression results in Table 4. The impact of information asymmetry on stock returns is positive (0.4 bps) but statistically insignificant. In the first column, we present the impact of information asymmetry on returns without controlling any other factors. In multivariate Fama–MacBeth regressions, the impact of PIN reduces down to 0.2 bps and is statistically insignificant for all cases¹. To this end, we fail to reject our first hypothesis which suggest that information asymmetry has no significant impact over the cross-section of stock returns². The results are robust to different lag length selections. We also compute the standard errors of the firm-level regressions using bootstrap sample sizes (Efron and Tibshirani 1986). The results survive in that setting as well.

Univariate portfolio analysis

We sort stocks into deciles based on PIN values estimated using the Duarte and Young (2009) factorization, for each month between March 2005 and April 2017. We form equally and value-weighted portfolios using the stocks in each decile and follow the returns of these portfolios for one, three and 6 months. In Table 5, we present the

¹Interestingly, contrary to the standard literature, we observe a significant and negative impact of BETA and BTM over future returns, and the positive impact of SIZE.

²We conduct a multicollinearity test by checking the variance inflation factors of the independent variables presented in Equation (8) in a contemporaneous ordinary least squares setting. Our results indicate that the standard errors for the coefficients are inflated less than factor of two. Therefore, we conclude that multicollinearity does not significantly affect the results presented in this manuscript.

Table 4. Fama–MacBeth regressions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
PIN	0.004 (0.94)	0.004 (1.00)	0.003 (0.85)	0.004 (0.98)	0.004 (0.96)	0.004 (0.90)	0.004 (0.94)	0.004 (1.05)	0.004 (−1.78)	0.003 (0.64)	0.002 (0.48)
BETA		− 0.002 (−1.22)									− 0.001 (−0.74)
SIZE			0.001 (2.79)								0.001 (2.88)
BTM				0.002 (1.92)							0.002 (1.53)
MOM					− 0.009 (−2.87)						− 0.008 (−2.74)
REV						0.025 (4.61)					0.030 (4.91)
ILLIQ							− 0.091 (−0.23)				− 0.222 (−0.53)
IVOL								− 0.166 (−3.01)			− 0.213 (−1.68)
MAX									− 0.029 (−1.78)		− 0.010 (−0.25)
PSOS										− 0.003 (−1.29)	− 0.005 (−1.70)

This table presents the results of Fama and MacBeth (1973) regressions. Each month, we regress the stock returns on previous months PIN estimates along with the control factors BETA, SIZE, BTM, MOM, REV, ILLIQ, MAX, IVOL and PSOS. Entries in the table are the time-series averages of the slope coefficients obtained from the cross-sectional regressions. Values in parenthesis presents t-statistics based on Newey and West (1987) standard errors with lag length equal to 1. The results are robust to different lag length selections such as {2, 3, 4, 5, 6, 12, 24 and 36}.

descriptive statistics for decile portfolios obtained with 1-month holding period³.

The portfolio statistics is set to be the equal-weighted averages of the stocks within that portfolio and the measures are calculated at the end of each month where the portfolios are formed. We observe that, by construction, the average PIN values increase from low-PIN decile to high-PIN decile. Other than PIN estimates, we do not observe a linear increase or decrease in any other measure. We observe the highest average monthly return in the ninth decile (average return 8%) and the second decile (average return is 4.5%). The

stocks in the ninth-decile also has the highest daily return volatility. On average, the stocks in the fifth-decile has the highest market capitalization. To this end, we can say that PIN-sorted portfolios in BIST does not have a monotonically increasing or decreasing feature which may suggest that information asymmetry does not have any systematic impact on stocks traded in BIST.

We examine whether the investors in BIST demand a premium for holding stocks that have high information asymmetry through examining the difference in returns of high-PIN and low-PIN portfolios ($R^{High} - R^{Low}$). In Table 6, we show the

Table 5. Descriptive statistics for PIN portfolios.

Deciles	Average PIN	Average Return	Median Return	Volatility	Skewness	SIZE	BETA	BTM	IVOL	MOM
Low-PIN	0.0000	0.015	−0.016	0.13	1.20	902	0.57	1.01	0.02	0.15
2	0.0004	0.045	0.013	0.13	1.78	1017	0.18	1.27	0.02	0.08
3	0.0038	0.027	0.001	0.10	1.19	1139	0.20	0.95	0.03	0.08
4	0.0472	0.017	−0.005	0.10	1.16	3515	0.36	1.24	0.02	0.09
5	0.1126	0.013	0.001	0.09	1.77	4938	0.48	0.99	0.02	0.10
6	0.1592	0.014	−0.006	0.10	1.16	2147	0.40	1.24	0.02	0.16
7	0.2083	0.026	−0.003	0.09	0.94	1723	0.30	1.18	0.02	0.13
8	0.2558	0.039	0.016	0.13	2.55	2150	0.36	1.05	0.02	0.12
9	0.3620	0.078	0.027	0.22	2.66	1296	0.31	1.12	0.03	0.18
High-PIN	0.5195	0.023	−0.002	0.09	0.83	1368	0.31	0.96	0.02	0.11

This table presents a descriptive statistics on the decile portfolios formed based on monthly PIN estimates. At the end of each month, we sort stocks based on PIN measures estimated on that month, and form decile portfolios based on those estimates. We follow the returns of stocks in portfolios next month and set the return of each portfolio to be the equal weighted average of the stock returns within those portfolios. The first column shows the average PIN estimate of each portfolio. The second column shows the average and the third column shows the median return, respectively. Fourth column shows the historical volatility of the monthly portfolio returns whereas the fifth column shows the skewness of returns. From the sixth column to the tenth column, we present the portfolios' SIZE, BETA, BTM, IVOL, MOM values, respectively.

³We observe similar values for three and 6-month holding periods. We do not report these tables to save space.

Table 6. Descriptive statistics for PIN portfolios.

Decile	Equally Weighted Portfolios		Value Weighted Portfolios	
	Average Return (%)	Three Factor alpha	Average Return (%)	Three Factor alpha
Panel A: Holding Period = 1 month				
Low-PIN	0.019	0.011	0.014	0.007
2	0.017	0.008	0.015	0.011
3	0.014	0.005	0.016	0.008
4	0.015	0.005	0.007	0.000
5	0.012	0.003	0.016	0.007
6	0.016	0.005	0.012	0.001
7	0.018	0.008	0.029	0.037
8	0.014	0.004	0.016	0.005
9	0.013	0.004	0.008	-0.003
High-PIN	0.014	0.006	0.009	-0.001
High-Low	-0.004	-0.005	-0.004	-0.008
t-statistics	(-1.16)	(-1.20)	(-0.73)	(-1.19)
Panel B: Holding Period = 3 months				
Low-PIN	0.018	0.009	0.018	0.013
2	0.016	0.007	0.007	-0.001
3	0.026	0.017	0.042	0.057
4	0.014	0.004	0.013	0.005
5	0.017	0.006	0.015	0.004
6	0.014	0.002	0.010	-0.001
7	0.015	0.004	0.014	0.001
8	0.014	0.005	0.027	0.023
9	0.014	0.004	0.010	-0.001
High-PIN	0.014	0.004	0.004	-0.006
High-Low	-0.004	-0.005	-0.014	-0.019
t-statistics	(-1.14)	(-1.29)	(-2.06)	(-2.73)
Panel C: Holding Period = 6 months				
Low-PIN	0.012	0.001	0.010	0.002
2	0.013	0.005	0.017	0.017
3	0.016	0.006	0.011	0.003
4	0.016	0.007	0.010	0.002
5	0.015	0.005	0.012	0.003
6	0.012	0.005	0.018	0.018
7	0.011	0.001	0.007	-0.003
8	0.016	0.006	0.011	0.002
9	0.013	0.004	0.014	0.004
High-PIN	0.019	0.010	0.016	0.006
High-Low	0.007	0.009	0.006	0.004
t-statistics	(1.85)	(2.22)	(0.80)	(0.49)

In this table, we present the equally and value-weighted returns of decile portfolios. At the end of each month, we sort our stocks into decile portfolios based on monthly PIN estimates obtained using Duarte and Young (2009) factorization. We follow the portfolio returns for one, three and 6 months. Panel A, B and C present the findings for one, three and 6 months holding period, respectively. The three factor α is obtained through regressing the portfolio returns on the market return (BIST100), and size (SMB) and profitability factors (HML) of Fama and French (1993). High-Low represents the zero investment PIN portfolio which mimics the returns of an investor who has a long position in High-PIN portfolio and a short position of equal size in the Low-PIN portfolio. t-statistics are calculated with Newey and West (1987) adjusted standard errors.

time-series analysis of zero investment PIN portfolios.

In Table 6, Panel A and B, we observe that the average returns on the zero investment PIN portfolios are negative for one-month and three-month holding periods, respectively. This finding contradicts with the theoretical expectations of Easley and O'hara (2004) and the empirical findings of Easley, Hvidkjaer, and O'Hara (2002), Easley, Hvidkjaer, and O'Hara (2010) and Lai, Ng, and Zhang (2014). However, we fail to reject the hypothesis that the returns of zero investment portfolios are, on average, different from zero. That is why the importance on the sign of the coefficient largely

disappears. In Table 6, Panel C, we observe that only the three-factor α for the equally weighted portfolio is positive and statistically significant. To this end, we can say that the 6-month premium for zero investment PIN portfolios are 90 bps, which is economically insignificant given the average monthly return is 1% in our sample.

In line with our regression analysis, our univariate portfolio analysis suggests that the information asymmetry proxied by PIN does not have any systematic impact on stock returns in BIST. We fail to reject the hypothesis that the return premium of zero investment PIN portfolios is equal to zero. In line with the expectations of Hughes, Liu, and Liu

Table 7. Characteristics and returns of PIN-size portfolios.

	Low	2	High	DIF
Panel A: SIZE				
Small	2.383	2.532	2.459	0.077
2	3.191	3.402	3.271	0.080
3	3.717	3.986	3.828	0.111
4	4.189	4.486	4.326	0.137
5	4.652	5.027	4.859	0.207
6	5.134	5.593	5.409	0.275
7	5.687	6.180	6.000	0.313
8	6.263	6.774	6.575	0.312
9	6.926	7.638	7.387	0.461
Large	8.915	9.445	9.218	0.303
Panel B: PIN				
Small	0.004	0.117	0.374	0.369
2	0.004	0.116	0.373	0.369
3	0.004	0.119	0.372	0.368
4	0.004	0.120	0.371	0.367
5	0.004	0.121	0.364	0.359
6	0.004	0.121	0.359	0.355
7	0.004	0.122	0.365	0.360
8	0.004	0.121	0.361	0.357
9	0.004	0.125	0.355	0.351
Large	0.005	0.128	0.350	0.345
Panel B: RETURN				
Small	0.020	0.022	0.013	-0.008
2	0.027	0.011	0.015	-0.013
3	0.012	0.016	0.010	-0.002
4	0.016	0.011	0.010	-0.006
5	0.015	0.019	0.019	0.004
6	0.014	0.018	0.014	0.000
7	0.016	0.016	0.013	-0.004
8	0.013	0.012	0.017	0.004
9	0.011	0.011	0.012	0.001
Large	0.017	0.020	0.014	-0.003

This table presents the characteristics of PIN-Size portfolios. At the beginning of each month, we divide the sample into size deciles based on the end-of-month market capitalization of previous month. Then, we divide each decile into three PIN groups, based on previous month's PIN estimates. Panel A contains the Turkish Lira denominated time-series averages of the mean market capitalization for each portfolio. In panel B, we present the time-series averages of the equally weighted means of PIN for each portfolio. Panel C represents the time-series average of average monthly returns in each portfolio. PIN-Size portfolios are constructed based on previous months market capitalization and PIN estimates. DIF is the average difference between High and Low-PIN portfolios. $t(DIF)$ is the t-statistics of DIF.

(2007), we show that the risk due to information asymmetry can be eliminated entirely within a one-month or three-month holding period, through diversification. For the six-month period, we show that the economic meaning of the premium due to information asymmetry largely disappears.

To a large extent, our univariate analysis show that risk due to information asymmetry is firm-specific for stocks that are traded in BIST. Lastly, in line with Easley, Hvidkjaer, and O'Hara (2010), we examine whether the firm size plays any role in explaining the impact of information asymmetry on stocks returns through multivariate portfolio analysis.

Multivariate portfolio analysis

We follow a similar methodology proposed in Easley, Hvidkjaer, and O'Hara (2010) to examine the impact of informed trading on securities returns.

We sort stocks with respect to their size at the beginning of each month and form 10 size portfolios. Then, again for each month, we divide the portfolios into three with respect to previous month's PIN estimates. This yields us the time series of 30 different portfolio returns. We follow the portfolio returns for one, three and 6 months. Then, we subtract the low-PIN portfolio returns from high-PIN portfolio returns in each size decile and form 10 zero investment PIN-Size portfolios (R_p $p = 1 \dots 10$). This gives us the mimicking portfolio returns of an investor who has a long position in high-PIN stocks and a short position with equal size in low-PIN stocks, controlling for size. The descriptive statistics on zero investment PIN-Size portfolios are provided in Table 7.

In Table 7 Panel A, we observe that for each size decile, the difference between the market value of high-PIN portfolios and low-PIN portfolios

monotonically increase, except for the last size decile. The highest size stocks are grouped in mid-PIN portfolios which is not expected. In addition, in Panel B, we observe that the difference in PIN values decrease as we move from small to big decile. Panel A and B indicate that there is no meaningful multivariate relationship between information asymmetry and size in BIST⁴. Lastly, we share the results on Hypothesis 3 on Panel C. We fail to reject that the difference between high-PIN and low-PIN returns in each size decile is equal to zero. This finding implies that under the control of market capitalization, the investors who invest in stocks that have high information asymmetry do not obtain premium.

Finally, we examine the systematic impact of information asymmetry on stock returns through Equation (9). Akdeniz, Altay-Salih, and Aydogan (2000) show that the size and profitability factors of Fama and French (1993) correspond to a systematic risk in BIST. The results on Fama and French (1993) regressions provided in Table 8, below.

In line with our previous findings, results presented in Table 8 indicate that information asymmetry does not provide a systematic return premium for stocks that are trading in BIST. The coefficient estimates for the constant terms, $\hat{\vartheta}_p^1$ are mostly negative and insignificant. The values range between -1 bps and 0.5 bps. We also fail

to reject the hypothesis that the constant terms are jointly equal to zero (p value(GRS) = 0.12). We then conclude that zero investment PIN-Size portfolios do not bring any premium after controlling for other systematic factors such as market return, size and profitability. With these results, our study provides strong empirical support to the theoretical expectations of Hughes, Liu, and Liu (2007). We show that information risk is idiosyncratic in nature and can be eliminated with diversification. Our results are robust to any possible bias due to trade classification, floating-point exception, boundary solutions and systematic order flow shocks.

VI. Conclusion

Standard asset pricing models ignore the impact of information asymmetry on equity returns. In the finance literature, significant number of studies focus on asset pricing anomalies that standard models cannot explain. Previous work on the role of information asymmetry provide conflicting results.

The rational expectations model of Easley and O'hara (2004) expects uninformed investors demand premium for holding stocks that have high information asymmetry. This is due to the fact that uninformed investors are unable to adjust their portfolios quickly to the new information

Table 8. Fama-French three factor regressions.

	Zero Investment PIN-Size Portfolios									
	Small	2	3	4	5	6	7	8	9	Large
$\hat{\vartheta}_p^1$	-0.004	-0.009	0.000	-0.008	0.005	-0.003	-0.004	0.005	-0.002	-0.010
	(-0.58)	(-1.55)	(-0.09)	(-1.61)	(1.00)	(-0.50)	(-0.82)	(1.31)	(-0.46)	(-1.83)
$\hat{\vartheta}_p^2$	-0.241	-0.164	0.287	-0.037	-0.045	-0.257	-0.083	-0.090	0.050	0.026
	-1.24	-1.12	2.29	-0.30	-0.33	-1.98	-0.72	-0.85	0.48	0.19
$\hat{\vartheta}_p^3$	(-0.36)	(-0.27)	(-0.33)	(0.00)	(-0.16)	(0.19)	(0.00)	(-0.05)	(0.19)	(0.47)
	-1.48	-1.45	-2.09	0.01	-0.89	1.17	0.02	-0.39	1.43	2.80
$\hat{\vartheta}_p^4$	0.077	-0.065	0.090	0.126	0.025	0.116	0.011	-0.029	0.059	0.115
	(0.75)	(-0.83)	(1.35)	(1.89)	(0.34)	(1.68)	(0.18)	(-0.51)	(1.07)	(1.63)
F-stat	2.21	2.38	2.31	1.54	0.47	3.58	0.24	0.56	1.97	5.76
Adj. R ²	0.02	0.03	0.03	0.01	-0.01	0.05	-0.02	-0.01	0.02	0.09
GRS Test Statistics	1.59									
p-value(GRS)	0.12									

We run the following regression on the time series of monthly returns of ten zero investment PIN-Size portfolios. PIN estimates are obtained through DY algorithm.

$$R_p = \hat{\vartheta}_p^1 + \hat{\vartheta}_p^2 R_m + \hat{\vartheta}_p^3 SMB + \hat{\vartheta}_p^4 HML$$

+ ϵ_p . The sample period is between March 2005–April 2017. We provide the coefficient estimates and their respective t-statistics. For goodness of fit, we provide adjusted R^2 measures along with F statistics. Last two rows provide the Gibbons, Ross, and Shanken (1989) test statistics for the null hypothesis that $\hat{\vartheta}^1 = 0$.

⁴As a robustness check, we reverse the sorting order when forming portfolios. That is, we sort our stocks first into three PIN groups, then we divide each PIN group into size deciles. Our results persist in this setting as well.

arrive at the market. On the contrary, Hughes, Liu, and Liu (2007) in a market with sufficient number of securities, only the market-wise information asymmetry will be priced and therefore firm-specific information risk is idiosyncratic in nature.

Numerous empirical studies have been conducted on hypotheses presented by these two opposing views. Easley, Hvidkjaer, and O'Hara (2002) and Easley, Hvidkjaer, and O'Hara (2010) find that the information asymmetry proxied by PIN is a systematic risk factor for U.S. markets. On the contrary, Duarte and Young (2009) and Lai, Ng, and Zhang (2014) questions the systematic impact of PIN factor.

Recently, the studies that examine the role of information asymmetry in BIST significantly increased. For instance, Aktas (2016) shows that the short sellers have higher price impact for BIST30 firms between 2008–2009. Similarly, Tiniç and Savaser (2018) examine the information asymmetry between foreign and domestic investors and underline the informational advantage of domestic investors over the foreigners. Simsir and Simsek (2018) document an abnormal volume of trades right before an announcement made public in public disclosure platform. This study is first in exploring the systematic impact of informed trading in equity returns in BIST.

We obtain the PIN estimates for all stocks that are traded in BIST between March 2005 and April 2017. To estimate the PIN, we use the factorization provided by Duarte and Young (2009) and a grid-search algorithm similar to Yan and Zhang (2012). We examine the relationship between returns and PIN in a Fama and MacBeth (1973) framework, along with univariate and multivariate portfolio analysis.

Our results are in line with the theoretical expectations of Hughes, Liu, and Liu (2007) and the empirical findings of Lai, Ng, and Zhang (2014). We find that the information asymmetry proxied by PIN, is an idiosyncratic risk and therefore it can be eliminated with portfolio diversification. We document a statistically insignificant relationship between PIN and future returns. With univariate and multivariate portfolio analysis, we mimic the portfolio of investors that has long position

in high-PIN stocks and a short position of equal size in low-PIN stocks. We show that investors do not obtain any premium for investing in these zero investment PIN portfolios. Our results do not depend on trade classification algorithms, different factorizations in estimating PIN. Similarly, our estimates are not effected by any bias due to floating-point exception, boundary solutions and systematic order flow shocks.

Acknowledgments

We thank Nuray Guner, Levent Akdeniz, Tanseli Savaser, Basak Tanyeri, Oguz Ersan and participants of Turkish Finance Workshop at Middle East Technical University, 18th Workshop on Quantitative Finance at Universita degli Studi di Milano, Summer School on Market Microstructure at Universita della Svizzera Italiana, Workshop on Market Microstructure and Behavioural Finance at Istanbul Technical University for helpful comments and suggestions. Authors gratefully acknowledge the financial support provided by the Scientific and Technological Research Council of Turkey (TUBITAK), Grant Number: 116K335. This research is conducted as a part of Tiniç's Ph.D. Thesis at Bilkent University, Faculty of Business Administration.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the The Scientific and Technological Research Council of Turkey (TUBITAK) [116K335].

References

- Akdeniz, L., A. Altay-Salih, and K. Aydogan. 2000. "A Cross-section of Expected Stock Returns on the Istanbul Stock Exchange." *Russian & East European Finance and Trade* 36 (5): 6–26.
- Aktas, O. U. 2016. "Three Essays on the Microstructure of the BIST." PhD thesis, Concordia University.
- Aktas, O. U., and L. Kryzanowski. 2014. "Trade Classification Accuracy for the Bist." *Journal of International Financial Markets, Institutions and Money* 33: 259–282.
- Alkan, U., and B. Guner. 2018. "Preferences for Lottery Stocks at Borsa Istanbul." *Journal of International Financial Markets, Institutions and Money* 55: 211–223.

- Amihud, Y. 2002. "Illiquidity and Stock Returns: Cross-section and Time-series Effects." *Journal of Financial Markets* 5 (1): 31–56.
- Atilgan, Y., K. O. Demirtas, and A. D. Gunaydin. 2016. "Liquidity and Equity Returns in Borsa Istanbul." *Applied Economics* 48 (52): 5075–5092.
- Celik, D., and M. Tiniç. 2018. "InfoTrad: An R Package for Estimating the Probability of Informed Trading." *The R Journal* 10 (1): 31–42.
- Duarte, J., and L. Young. 2009. "Why Is Pin Priced?" *Journal of Financial Economics* 91 (2): 119–138. doi:10.1016/j.jfineco.2007.10.008.
- Easley, D., and M. O'hara. 2004. "Information and the Cost of Capital." *Journal of Finance* 59 (4): 1553–1583.
- Easley, D., N. M. Kiefer, M. O'Hara, and J. B. Paperman. 1996. "Liquidity, Information, and Infrequently Traded Stocks." *The Journal of Finance* 51 (4): 1405–1436. doi:10.1111/j.1540-6261.1996.tb04074.x.
- Easley, D., S. Hvidkjaer, and M. O'Hara. 2002. "Is Information Risk a Determinant of Asset Returns?." *The Journal of Finance* 57 (5): 2185–2221. doi:10.1111/1540-6261.00493.
- Easley, D., S. Hvidkjaer, and M. O'Hara. 2010. "Factoring Information into Returns." *Journal of Financial and Quantitative Analysis* 45: 293–309. doi:10.1017/S0022109010000074.
- Efron, B., and R. Tibshirani. 1986. "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy." *Statistical Science* 1: 54–75.
- Ersan, O., and A. Alici. 2016. "An Unbiased Computation Methodology for Estimating the Probability of Informed Trading (pin)." *Journal of International Financial Markets, Institutions and Money* 43: 74–94. doi:10.1016/j.intfin.2016.04.001.
- Fama, E. F., and J. D. MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* 81 (3): 607–636.
- Fama, E. F., and K. R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33 (1): 3–56.
- Gan, Q., W. C. Wei, and D. Johnstone. 2015. "A Faster Estimation Method for the Probability of Informed Trading Using Hierarchical Agglomerative Clustering." *Quantitative Finance* 15 (11). doi:10.1080/14697688.2015.1023336.
- Gibbons, M. R., S. A. Ross, and J. Shanken. 1989. "A Test of the Efficiency of A Given Portfolio." *Econometrica: Journal of the Econometric Society* 57: 1121–1152.
- Hughes, J. S., J. Liu, and J. Liu. 2007. "Information Asymmetry, Diversification, and Cost of Capital." *The Accounting Review* 82 (3): 705–729.
- Jegadeesh, N. 1990. "Evidence of Predictable Behavior of Security Returns." *The Journal of Finance* 45 (3): 881–898.
- Lai, S., L. Ng, and B. Zhang. 2014. "Does Pin Affect Equity Prices around the World?" *Journal of Financial Economics* 114 (1): 178–195.
- Lin, H.-W.-W., and W.-C. Ke. 2011. "A Computing Bias in Estimating the Probability of Informed Trading." *Journal of Financial Markets* 14 (4): 625–640. doi:10.1016/j.finmar.2011.03.001.
- Newey, W. K., and K. D. West. 1987. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica: Journal of the Econometric Society* 55: 703–708.
- Simsir, S. A., and K. Simsek. 2018. "Measuring the Market Impact of Private Information: The Curious Case of Corporate Announcements in Turkey." Working Paper, Sabanci University, Sabanci School of Management.
- Tiniç, M., and T. Savaser. 2018. "Political Turmoil and the Role of Foreign Orders on Equity Prices." Working Paper, Bilkent University, Faculty of Business Administration.
- Yan, Y., and S. Zhang. 2012. "An Improved Estimation Method and Empirical Properties of the Probability of Informed Trading." *Journal of Banking & Finance* 36 (2): 454–467. doi:10.1016/j.jbankfin.2011.08.003.

Appendix

Variable definitions

SIZE: Firm size for a given month is taken to be the natural logarithm of the firms' market value, that is, the end of month price times the number of shares outstanding.

BOOK TO MARKET RATIO: Book to market ratio for a firm is taken to be the ratio of the firms end of month total equity value to its market value.

BETA: We estimate the market risk of a given stock i , on a given month t , from the market model given below;

$$R_{id} = \alpha_i + \beta_i R_{md} + e_{id} \quad d = 1 \dots D_t \quad (10)$$

where D_t is the number of trading days on month t . R_{id} is the daily return on the stock i on a given day d . Similarly, we take R_{md} to be the return of the BIST100 index of a given day d . We estimate Equation (10) for each stock using daily returns within a month.

ILLIQUIDITY: We measure stock illiquidity for each month, t , with the ratio of the absolute monthly stock return to its trading volume, similar to Amihud (2002).

$$ILLIQ_{it} = |R_{it}| / VOL_{it} \quad (11)$$

Idiosyncratic volatility: We measure the idiosyncratic risk of a stock from the standard deviation of the daily residuals, e_{id} , presented in (10). That is,

$$IVOL_{it} = \sqrt{\text{var}(e_{id})} \quad (12)$$

REVERSAL: We define the reversal variable for each stock in month t as the return on the stock over previous month, similar to Jegadeesh (1990).

MOMENTUM: We define the momentum variable for each stock in month t as the cumulative percentage return over the last 6 months. That is,

$$MOM_{it} = \frac{P_{it-1} - P_{it-7}}{P_{it-7}} \quad (13)$$

where P_{it} is the price of the stock at the end of month t .

MAX: We define maximum daily return within a month as:

$$MAX_{it} = \max(R_{id}), \quad d=1 \dots D_t$$

where D_t is the number of trading days on month t , R_{id} is the daily return on the stock i on a given day d