Running head: INFORMATION RISK PRICED

RĪGAS EKONOMIKAS AUGSTSKOLA
STOCKHOLM SCHOOL OF ECONOMICS IN RIGA

Is Information Risk Priced in the Baltic Stock Markets?

Author: Saulius Nižinskas
Supervisor: Alminas Žaldokas
Stockholm School of Economics in Riga

Bachelor thesis

May 2009
Riga

COPYRIGHT DECLARATION

Name of the author
Saulius Nižinskas

Title of the Thesis

Is Information Risk Priced in the Baltic Stock Markets?

Signed

Saulius Nižinskas ______________________

Date

……………………………………….

Abstract

Recent years have seen an increasing popularity and application of the concept of the probability of information-based trading (PIN) proposed by Easley, Hvidkjaer & O'Hara (EHO) (2002). Authors provide a preposition that uninformed investors because of information asymmetry require a compensation for holding stocks that generate more private information events. Thus, this research investigates whether information risk (as proxied by the PIN) determines stock returns in the Baltic equity markets. The sequential trade model provides PIN scores which reveal that the lowest levels of information asymmetry are present in Tallinn Stock Exchange (TSE), followed by Vilnius Stock Exchange (VSE) and Riga Stock Exchange (RSE). The fraction of informed traders, the frequency of information events, and company characteristics (e.g. corporate governance level) are possible main factors for differences in the PIN across Baltic stock exchanges. The significance of information risk for stock returns is tested incorporating the PIN as additional variable to Fama and French (1992) three factor model. It is found that the PIN has a weak effect on excess returns in VSE: 10% increase in PIN leads to 3.8% increase in excess returns. The PIN is found to have no effect in RSE, TSE and pooled data sample. The possible demand factors (e.g. activity of small traders, perception of risk) or supply factors (e.g. company specific characteristics) are determinants of these differences. The analysis also reveals that there is an inverse market risk return trade-off, large and "value" stocks outperform small and "growth" stocks, respectively.

Keywords: Information-based trading, information risk, information asymmetry, PIN, asset pricing, Baltic stock exchanges.

Table of Content

1 Introduction

The research on asset-pricing in several past decades was largely based on traditional asset pricing theories that focus on the role of market and other aggregative risks. Only recent studies have seen an increasing attention towards market microstructure factors and their explanatory power for asset returns. Easley, Hvidkjaer & O'Hara (EHO) (2002) provide a proposition that the process of how prices incorporate new information and become efficient is not static but dynamic. Thus, how prices become efficient cannot be separated from the evolution of trading prices. This creates a probability that some investors at a certain time are better informed than others or that different investors interpret information differently. This idea further leads to the fact that, in equilibrium, informed and uninformed investors hold different market portfolios. The portfolios of uninformed market participants include more stocks that convey "bad news" to the market while the portfolios of informed investors consist of more "good news" stocks. Since the market is fully aware of these information imbalances, in equilibrium, a uniformed investor must be rewarded for the risk of holding a stock with private information events; therefore, this information risk should be priced in asset returns as well. EHO (2002) generate a measure of the probability based trading, PIN, and show that information does affect asset returns while incorporating it into Fama and French three factor asset pricing model for NYSE-listed stocks.

Vilnius, Riga and Tallinn stock exchanges can be considered as relatively young and emergent equity markets. Currently they belong to NASDAQ OMX group where the order-driven trading mechanism is applied and multiple orders can be submitted (limit order, market order, etc.). The transaction orders are entered and matched automatically without a direct involvement of a broker. Stock exchanges also use daily price limit bounds. These characteristics of Baltic stock markets make it interesting to test whether any conclusions about information risk and its explanatory power for asset returns can be drawn in Vilnius Stock Exchange (VSE), Riga Stock Exchange (RSE), Tallinn Stock Exchange (TSE) and overall Baltic market.

This paper aims at providing additional arguments about emergent markets, and more specifically, filling a gap of investigating information risk and asset returns in the context of Baltic equity markets. Therefore, it is believed that this research will give a ground for a further investigation and interest in the information-based trading and asset pricing in the Baltic stock exchanges. Moreover, this paper also contributes to previous research by providing extra empirical evidence on information asymmetry, its explanatory power for stock returns and the risk-return trade-off. From a practical point of view, it is expected that this paper will give another stimulus for Baltic policy makers to ensure the equal access to information for market participants. In addition, the paper presents empirical evidence of what drives stock returns in the Baltic stock markets which should be highly relevant for ordinary traders.

Hence, the research question is stated: **“To what extent the probability of information-based trading affects stock returns in the Baltic equity markets?”**

In order to address this question, the following hypotheses are analyzed and tested:

H1: The probability of information-based trading is a significant factor of stock returns in VSE;

H2: The probability of information-based trading is a significant factor of stock returns in RSE;

H3: The probability of information-based trading is a significant factor of stock returns in TSE;

H4: The probability of information-based trading is a significant factor of stock returns in all Baltic stock exchanges (pooled data).

In order to test these hypotheses, the paper rests on the concept of the probability of information-based trading and the market microstructure model developed by EHO (2002). The study utilizes the daily trading data gathered for Baltic listed companies for 2004Q4 – 2008Q3 time period. The PIN is applied as a proxy for information risk. It is afterwards incorporated to Fama and French three factor model. The significance of PIN is tested running OLS regressions and afterwards averaging time-series coefficients under Fama-MacBeth two-step methodology (Fama & MacBeth, 1973) and Litzenberger and Ramaswamy (1979) correction technique. In order to test first three hypotheses, separate OLS regressions are run taking into account only stocks that belong to a particular market, while the fourth hypothesis is tested running regressions with pooled data set from all Baltic markets.

The paper is structured as follows: *Section 2* provides an insight to relevant literature on the topic and related issues; *Section 3* describes the theoretical framework and methodology; *Section 4* presents the data gathered, while *Section 5* overviews and discuss main findings, and *Section 6* concludes the paper.

2 Literature Review

This section provides a brief overview about market microstructure models and previous research. The section is split into four main parts. The first part is dedicated for microstructure models. Despite the fact that the concept of the PIN is relative new concept, it has received a fast and great recognition. Therefore, studies that formed a background of the probability of information-based trading concept are discussed. Simultaneously, the major empirical findings, that were discovered applying the PIN, are outlined. Furthermore, papers that specifically relate information risk and asset pricing are overviewed. Last but not least, studies that investigated Baltic markets and that are relevant for this study are summarized.

2.1 Market Microstructure Models

Market microstructure models can be classified by several dimensions: type of orders, sequence of moves, price setting rules, and competitive vs. strategic structures. Most models share a common characteristic that the price of a stock adjusts instantly to new public information but only gradually to private information. This gradualism takes place because there is noisy asset supply and/or strategic behaviour of informed traders (Brunnermeier, 2001).

The PIN model developed by EHO (2002) can be considered as a microstructure model in which participants of a market draw their inferences about the true value of an asset according to the trading activity flow. In many sequential trade models with multiple trading rounds and competitive market makers the process of learning is observed (see Brunnermeier (2001), for more explicit explanations of models see Glosten and Milgrom (1985), Kyle (1985)). In such models market participants constantly update their beliefs about the private information based on trade imbalances (e.g. buy and sell trades) and set trading prices. Over a day, because of the process of trading, learning and price setting, asset prices converge to full information levels as the private information is revealed due to activity of informed market participants. The exclusive feature of EHO (2002) model is that it introduces a concept of "event uncertainty". This characteristic allows informed traders do not engage in a trade if there was no information event happening. Since this is much more precise approximation of real events in a market, EHO (2002) sequential trade model will be applied in this research.

2.2 PIN Derivation and Application

The set of influential recent papers about market microstructure and, particularly, the information risk has been written over past two decades (see Easley, Kiefer & O'Hara (1996, 1997a, 1997b), Easley, Kiefer, O'Hara & Paperman (1996), and Easley, O'Hara & Paperman (1998)). The authors popularize the concept of the probability of information-based trading.

Based on the sequential market microstructure model and trade order flow data, the measure PIN, that proxies information asymmetry, is derived. Applying the PIN authors address number of issues: there is a significant difference between information content of orders executed in the New York Stock Exchange and the Cincinnati Stock Exchange; large and small trades have a different information content and they are almost equally informative; high volume stocks, on average, yield the higher probability of information-based trading which also has a significant economic importance for spreads, and verify if financial analysts engage in informed trading.

These papers have recently received much attention, with most of it concentrated on the PIN application for different issues in corporate finance, investments and market microstructure. For instance, Heidle & Huang (2002) investigate market structure and different trading mechanisms applying the PIN measure. Their evidence suggests that the probability of encountering trades by informed market participants is greater in dealer markets than in auction markets. The PIN is also found to be positively associated with higher bid-ask spreads. Brown et al. (2004) investigate what determines the information asymmetry and choose the PIN to capture it. Economic theory implies that information asymmetry can be diminished by greater disclosure. The authors prove that the frequency of conference calls is a highly significant determinant of subsequent level of information asymmetry. Since the cost of capital is increasing with information asymmetry, Brown et al. (2004) results imply that firms which make more conference calls have a lower cost of capital. This is consistent with Duarte et al. (2006) who investigate firms' information environments using the PIN. They demonstrate that the introduction of Regulation Fair Disclosure significantly influenced NASDAQ firms' cost of capital while there was no effect on NYSE/AMEX firms. Therefore, they partially prove that changes in firms' cost of capital are negatively related with the PIN. This issue was also investigated by EHO (2005b) who prove that investors demand higher return for holding stocks with private information. In addition, Agarwal & O'Hara (2006) demonstrate that firm's market leverage is affected by information risk: higher information asymmetry leads to a higher usage of debt as a part of capital structure. This is further supported by Bharath et al. (2008).

Even though the PIN has not been yet applied for Baltic equity markets, it was already employed in investigating information-based trading in another Central and Eastern European market, the Czech Republic. Nemecek (1997) reports that the PIN is lower for more actively traded stocks. This is in line with Easley, Kiefer & O'Hara (1996). Hanousek & Podpiera (2002) adds that information-based trading in Prague Stock Exchange is much more apparent than in other well-established equity markets and partially prove that informed and insider trading is widespread in the emerging financial markets of transition countries. Further, Kopriva (2008) tries to identify sources of the PIN and illustrates that market makers have very strong position in the Prague Stock Exchange. Findings of Kopriva (2008) and Hanousek & Podpiera (2002) imply that current regulation of the market should be strengthened.

Despite the fast recognition and application of PIN concept, some researchers identify issues that might cause bias in the PIN estimation. Boehmer et al. (2007) address the issue of a trade misclassification. The estimation of the PIN requires knowing a number of buys and sells per day. Therefore, Lee-Ready (1991) algorithm, which is usually applied in the PIN estimation to find out the trade direction, might induce a trade misclassification. According to Boehmer et al. (2007), this misclassification leads to downward-biased PIN estimates while the magnitude of the bias is related to security's trading intensity. In addition, Yan and Zhang (2006) notes that another bias arises when a maximization procedure of a likelihood function is ran. The problem is that numerical maximization solutions often fall on a corner of the parameter space (boundary solutions). This implies that at certain time periods for some stocks information events happen each day or do not happen at all. The authors propose a method to overcome this bias by exploring the parameter space more effectively. Moreover, Aktas et al. (2006) question not econometric but empirical properties. The authors show that the behaviour of the PIN is contradictive with clear evidence of information leakages around major corporate events.

Thus, even though some mentioned findings do not directly relate to PIN and stock returns, which is the primary subject of this research, the rapidly increasing recognition and wide PIN application demonstrates that the concept of the information-based trading is topical. The conception of PIN is successfully applied in studies related to corporate finance, investment management, corporate governance, market microstructure, etc. which indicates that the PIN is versatile. However, one should be always careful about computational properties of PIN estimation and double check the economic rationale of parameters in order to obtain meaningful results.

2.3 Information and Asset Returns

When considering information effects and their explanatory power for asset returns, it is often referred to liquidity which is the possibility of buying and selling assets without incurring additional loss (Kohn, 1994). Still, in a case if such a loss takes place, and investors include and net it, in equilibrium, they expect to be rewarded additionally for holding such an asset. This idea was popularized in Amihud and Mendelson (1986). The authors proxy liquidity by bid-ask spread and prove that it is a significant determinant of stock prices. Their evidence was supported by other studies as well (see Amihud and Mendelson (1989), Brennan and Subrahmanyam (1996), Amihud, Mendelson and Lauterbach (1997) and Amihud (2000)). However, Eleswarapu and Reinganum (1993) and Chalmers and Kadlec (1998) demonstrate that investors do not require any premium. These inconsistencies were questioned by EHO (2002). The authors provide a proposition that investors demand an extra compensation not for exogenous illiquidity, that results in larger spreads, but for information risk (which emerge from uninformed traders having incomplete information compared with informed traders) which would also predetermine these larger spreads. Clearly, illiquidity and information risk are related subjects. However, they are not the same. The illiquidity usually arises because of some exogenous factors (interest rates, tax, limited competition) while information risk is related with dynamic market efficiency. EHO (2002) applies the PIN and show that it is correlated with illiquidity (expressed as bid-ask spreads, volume, turnover and variability of returns). This might imply that the PIN works as a proxy for illiquidity. However, authors demonstrate that illiquidity does not affect assets returns while the PIN is found to be a significant determinant: stocks with higher levels of the probability of information-based trading yield higher expected returns. For example, two stocks with PIN difference of 10% lead to 2.5% p. a. difference in expected returns. These empirical results are achieved with cross-sectional regressions for NYSE-listed stocks for 1983-1998. Furthermore, EHO (2005a) demonstrate that the size neutral zero-investment portfolios, long in stocks with high PIN and short in stocks with low PIN earn significant abnormal returns.

Information-based trading and asset returns have been investigated in other equity markets as well. Fuller et al. (2007) were interested if any similar conclusions can be drawn for NASDAQ-listed stocks. They find that certain stock characteristics and the PIN correlate differently for NASDAQ stocks than that of EHO (2002) NYSE stocks. Additionally, the PIN is found to be only weakly priced and has no effect on increasing excess returns. Furthermore, the PIN has been tested for emergent markets as well. Wong et al. (2008) demonstrate that uninformed Chinese investors are compensated for information risk they bear. In addition, Lu and Wong (2007) provide evidence that information risk is a significant determinant of stock prices in Taiwan Stock Exchange, where 10% increase in the PIN leads to from 4 to 7% increase in annual stock returns.

Still, some researchers doubt on whether the PIN reflects information risk systematically priced by investors. Mohanram and Rajgopal (2007) demonstrate that the PIN effect on stock prices is not robust to alternative specifications and time periods. Lambert et al. (2006) argue that information risk should not be priced because of the fact that large number of traders diversifies away such risk. This is further supported by Hughes et al. (2005) who note that information risk should be diversified or subsumed by other existing risk factors in the multi-factor asset-pricing model. Duarte and Young (2009) decomposes PIN into two components out of which one is related with asymmetric information and another one with illiquidity. Authors demonstrate that only latter affects stock returns.

In conclusion, previous research of asset pricing and information risk gives striking evidence that stock returns are driven by the probability of the information-based trading. It also shows that even though the PIN and illiquidity are related issues, their nature differs. Nevertheless, the critique on the PIN persists mainly centred to information risk as unsystematic, not robust to model specifications and PIN decomposition. Therefore, it is expected that the results of this paper could be used as another argument of empirical properties of the PIN.

2.4 Research on the Baltic Equity Markets

To the best of my knowledge there is no research on information based trading in the Baltic context. Nevertheless, several papers provide some insight to the issue. Probably the closest research has been presented by Brannas et al. (2006). The authors hypothesize that information events that happened in the US and Russian stock markets should affect stock returns and, particularly, their volatility in the Baltic equity markets. The study was performed on country's stock market level. Therefore, they winnow out information content if it conveys bad or good news; however, inferences whether information is private or public are not made. The authors also disregard different types of traders (informed vs. uninformed). Applying

ARasMA-asQGARCH model, Brannas et al. (2006) prove that Latvian stock market indices are unaffected by news from New York and Moscow; Estonian stock market reacts stronger according to news from US than from Russia while VSE is impacted more by news from Moscow.

Hogfeldt et al. (2000) investigate stock returns in Central and Eastern Europe. Out of three Baltic states only Estonian market is researched. The authors apply Arbitrage Pricing Theory (APT) and choose different country specific explanatory variables. Because of scare and unreliable data fundamental risk and company-specific factors are not included. This is partially revised in Devyzis and Jankauskas (2004) who illustrate that small stocks yield higher returns; value stocks outperform growth stocks and there is a momentum effect. However, they find that liquidity is not priced.

To conclude, previous research on Baltic stock markets was only partially related to asset pricing and/or information risk. Therefore, it is expected that this paper will narrow this gap and foster the further investigation.

3 Methodology

This section presents the theoretical framework and methodology for the research. The basics of PIN and rational expectations equilibrium asset-pricing model are followed by the chosen approach of this paper.

3.1 Theoretical Framework

The first subsection covers the PIN measure of the probability of information-based trading, the central variable to this study. It is followed by the rationale why information asymmetry should influence asset returns.

3.1.1 PIN Model

The basic structure of the EHO model (2002) can be summarized as a sequential trade tree in Figure 1 (Appendix A). At the beginning of a day, there is α probability that an information event occurs and $(1 - \alpha)$ of no news. There is δ likelihood that the event conveys a bad signal to the market and $(1 - \delta)$ probability that the event conveys a good signal. Consequently, bad events happen with a probability of $\alpha\delta$ and good information events occur with a probability of $\alpha(1 - \delta)$. It is assumed that trades arrive during a day sequentially, following the Poison processes. Orders from informed traders arrive randomly at the rate μ only on "good news" or "bad news" days, orders from uninformed buyers arrive randomly at the rate ε_b, and orders from uninformed sellers arrive randomly at the rate ε_s. Informed traders buy if they perceive the event as a good signal and sell if the event conveys bad news. It is assumed that uninformed investors trade only for liquidity purposes. Moreover, it is supposed that these arrival rates are common knowledge to the market (but not for econometrician); therefore, market participants update their beliefs and set new prices according to them. Trades evolve and the price movements reflect changing beliefs of market participants. Thus, at the beginning of a day the probability of information-based trading can be summarized as:

$$PIN = \frac{\alpha\mu}{\alpha\mu+\varepsilon_b+\varepsilon_s} \qquad (3.1.1)$$

The PIN measure demonstrates the fraction of trades that are information-based. It also yields an economic intuition: the probability of information trades increases if there are more information events (α), more information traders (μ), and decreases with more activity of uninformed traders (ε_b and ε_s).

EHO (2002) proposes the likelihood function induced by this model:

$L\big((B,S)\big|\theta\big) =$

$$\alpha(1-\delta)\frac{(\mu+\varepsilon_b)^B(\varepsilon_s)^S e^{-(\mu+\varepsilon_b+\varepsilon_s)}}{B!S!}+\alpha\delta\frac{(\mu+\varepsilon_s)^S(\varepsilon_b)^B e^{-(\mu+\varepsilon_b+\varepsilon_s)}}{B!S!}+(1-\alpha)\frac{(\varepsilon_s)^S(\varepsilon_b)^B e^{-(\varepsilon_b+\varepsilon_s)}}{B!S!}$$

(3.1.2)

where *B* and *S* stands for total buy trades and total sell trades during a day, respectively, and $\theta = (\alpha, \delta, \mu, \varepsilon_b, \varepsilon_s)$ is the parameter vector. This likelihood function can be regarded as a mixture of three possible outcomes (no news, good news, bad news) with their weighted probabilities. Assuming sufficient independence conditions across days, the likelihood function for a period is:

$$V = L(\theta|M) = \prod_{i=1}^{N} L(\theta|B_i, S_i) \qquad (3.1.3)$$

where (B_i,S_i) is trade data for day $i=1,\ldots,N$ and $M=((B_1,S_1),\ldots,(B_N,S_N))$ is the data set. Maximizing (3.1.3) over θ parameter vector given the data set *M* provides us with the estimates for the parameter vector, from which we can calculate the PIN measure of information trading.

After dropping a constant term and rearranging, EHO (2005a) proposes a following factorization of the log likelihood function:

$$L((B_i,S_i)_{i=1}^{N}|\theta) = \sum_{i=1}^{N}[-\varepsilon_b-\varepsilon_s+M_i(\ln x_b+\ln x_s)+B_i\ln(\mu+\varepsilon_b)+S_i\ln(\mu+\varepsilon_s)]+$$

$$\sum_{i=1}^{N}\ln[\alpha(1-\delta)e^{-\mu}x_x^{S_i-M_i}x_b^{-M_i}+\alpha\delta e^{-\mu}x_b^{B_i-M_i}x_s^{-M_i}+(1-\alpha)x_s^{S_i-M_i}x_b^{B_i-M_i}] \qquad (3.1.4)$$

where $M_i = (min(B_i,S_i) + max(B_i,S_i))/2$, $x_s = \frac{\varepsilon_s}{\mu+\varepsilon_s}$, and $x_b = \frac{\varepsilon_b}{\mu+\varepsilon_b}$. According to EHO (2005a), this factorization increases computing efficiency and reduce truncation error.

Thus, by knowing only the total number of buy trades and sell trades per day, it is possible to run maximum likelihood procedure and get θ parameter vector estimates for certain period. Weighting the fraction of trades that are information-based provides the PIN, the measure for the information risk for each stock.

3.1.2 Information Asymmetry and Asset Pricing

In order to demonstrate why information asymmetry should affect asset prices Easley and O'Hara (2005b) proposes a simple rational expectations equilibrium asset-pricing model which derives the expected excess return on a single stock i:

$$E[v_i - p_i] = \frac{\delta \bar{x}_i}{\rho_i + (1 - \alpha_i) I_i \gamma_i + (1 - \mu_i) \alpha_i I_i \theta_i} \qquad (3.2.1)$$

where v is future value of stock i distributed $N(\bar{v}_i, \rho_i^{-1})$; p is the price today of stock i, x is per capita supply of stock i, I is the number of information signals distributed $N(v_i, \gamma_i^{-1})$; α is fraction of information signals that are private, μ is fraction of traders that receive private information, δ is coefficient of risk aversion, and θ is the precision of the uninformed traders' posterior distribution on the value of stock i. If traders are risk averse ($\delta > 0$) and the net stock supply is positive ($x > 0$), then these traders should be compensated for holding positive supply of stock i. Therefore, the future value of stock i should be higher than its price today. If there is perfect information ($\rho = \infty, \gamma = \infty$) available in the market, investors know the true value of stock. Thus, the stock becomes risk free and its price is equal to expected future value. In such equilibrium, all investors must hold the same portfolio of assets. This provides the main proposition of the model that if we take two identical stocks except that in one there is more private information, the stock with more public information will yield lower expected excess returns. This is because of the fact that when information events are private, uninformed investors cannot reveal it from price movements and, therefore, consider such stocks as being more risky.

In this case, uninformed traders are put into disadvantageous situation. Informed traders are able faster adjust their portfolio composition according to new information while uninformed traders end up always holding too much stocks with "bad news" or too little of stocks with "good news”. Uninformed investors could hold only money and in such way avoid this risk. However, from diversification point of view, their utility is higher while holding some risky stocks. This model demonstrates that the standard separation theory does not hold here since uninformed and informed traders perceive a different risk-return trade-off and because of that hold different portfolios. Private information provokes a new type of systematic risk that investors require compensation.

3.2 Approach

3.2.1 PIN Estimation

In this research, in order to estimate the PIN for stocks of VSE, RSE, and TSE, the modified Newton-Raphson method was applied using the SAS Non-Linear Programming procedure. The procedure code can be seen in Appendix B. The Newton-Raphson method is well-known algorithm in mathematics to find roots of equations. Therefore, here it is used to maximize the log likelihood function (3.1.4) that was suggested by Easley and O’Hara (2005a). The Newton-Raphson method starts the search of maximum value of the log likelihood function with initial values of parameters set randomly. The process stops when convergence criteria are satisfied. The maximization of log likelihood function in EHO (2002) was performed on yearly basis excluding stocks that did not have at least 60 trading days. Other studies ran maximum likelihood procedure on various time basis (monthly, quarterly, yearly) applying from 20 to 90 trading days as minimum liquidity requirement for a stock (see Brown et al. (2004), Vega (2006)). This research follows Duarte et al. (2006) and 1) calculates the PIN variable on quarterly basis 2) exclude stocks that have less than 20 trading days per quarter. According to Easley, Kiefer, and O’Hara (1996) these characteristics ensure the reasonably precise estimation of parameters and compliance with model’s stationarity assumptions.

3.2.2 Interrelationship of variables

The interrelationship analysis of information asymmetry and stock return will be performed on three levels. First, correlation matrix will be overview. Second, two portfolios will be formed based on high and low PIN scores. Finally, regression analysis will be performed to check if correlation implies causation.

In order to yield for some comparability with previous research on asset pricing, this research employs Fama and French (1992) three factor model. Two additional variables, size and book-to-market, that drive asset returns were discovered by the authors in their cross-sectional study. In addition, the research by EHO (2002) demonstrates that, after controlling for the Fama and French factors, private information (as proxied by the PIN) significantly affects asset returns. Therefore, this study also considers the influence of information risk and its impact on stock returns in the Baltic equity markets.

The application of model tested here involves the following cross-sectional regression:

$$R_{i,t} = \gamma_0 + \gamma_1(\hat{\beta}_{i,t}) + \gamma_2(PIN_{i,t-1}) + \gamma_3(SIZE_{i,t-1}) + \gamma_4(BM_{i,t-1}) + \eta_{i,t} \quad (4.2.1)$$

where $R_{i,t,}$ is an excess return on stock *i* at time *t*, *PIN* is a log of firm's lagged quarterly *PIN*, $\hat{\beta}$ is an equity beta estimate, *SIZE* is a log of firm's lagged capitalization in EUR, and *BM* is a log of company's quarterly lagged book equity value divided by lagged quarterly market equity value. The independents variables in the regression are lagged since econometrician cannot know more than ordinary trader at time *t*. In order to test first three hypotheses, the regression is run for each Baltic equity market separately, and to test fourth hypothesis for a pooled data set. Beta estimates are calculated as proposed by Campbell et al. (1997):

$$\beta_{im} = \frac{Cov[Z_i, Z_m]}{Var[Z_m]} \qquad (4.2.2)$$

where Z_m is the excess return on market portfolio and Z_i is the excess return on stock i. Beta estimates are calculated using monthly stock excess returns two years before the test quarter since such method provides the biggest data set. Stocks that do not have price data for past two years are excluded from the analysis. Alternatively to EHO (2002) this research does not sort stocks into portfolios since this leads to information loss (see Kim (1995), or Berglund & Knif (1999)), creates a risk of data snooping due to arbitrary sorting criteria (see Conrad et al. (2003)) or ruins significant relationships by the sorting procedure (see Berk (2000)).

The coefficients of cross-sectional regression are averaged applying Fama-MacBeth (1973) two-step methodology. Firstly, given T periods of data, (4.2.1) is estimated using OLS for each t, $t=1,...,T$, giving the T estimates of γ_{0t}, γ_{1t}, γ_{2t}, γ_{3t}, and γ_{4t}. Secondly, the times series of $\hat{\gamma}_{0t}$,$\hat{\gamma}_{1t}$,$\hat{\gamma}_{2t}$,$\hat{\gamma}_{3t}$ and $\hat{\gamma}_{4t}$ are analyzed. The interest of this research lies within $\hat{\gamma}_{2t}$, the coefficient for the PIN. Significant and positive value of $\hat{\gamma}_{2t}$ would imply that higher probability of information based trading translates to higher required return. Thus, the opposite would mean that hypotheses H1, H2, H3 and H4 are rejected. These implications are tested using simple t-test where t-statistic is defined as follows:

$$t(\bar{\hat{\gamma}}_j) = \frac{\bar{\hat{\gamma}}_j}{\bar{\hat{\sigma}}_{\gamma_j}}, \qquad (4.2.2)$$

where

$$\bar{\hat{\gamma}}_j = \sum_{t=1}^{T} \omega_{jt}\hat{\gamma}_{jt}, \qquad (4.2.3)$$

and

$$\bar{\hat{\sigma}}^2_{\gamma_j} = \sum_{t=1}^{T} \omega^2_{jt}(\hat{\gamma}_{jt} - \bar{\hat{\gamma}}_j)^2. \qquad (4.4.4)$$

Fama and MacBeth (1973) assume stationary distribution and weight parameter coefficients and variances with $\omega_{jt} = 1/T$ for all j and t. Since this methodology is not efficient under finite sample size and time-varying volatility this research follows correction method by

Litzenberger and Ramaswamy (1979). The correction technique avoids possible complication of errors-in-variables (that might emerge because beta estimates rather than true betas and not portfolio betas are applied), assumes heteroscedasticity and suggests using

$$\omega_{jt} = \frac{1/\sigma^2_{\hat{\gamma}_{jt}}}{\sum_t 1/\sigma^2_{\hat{\gamma}_{jt}}} \qquad (4.4.5)$$

where $\sigma^2_{\hat{\gamma}_{jt}}$ is the variance of $\hat{\gamma}_{jt}$. Litzenberger and Ramaswamy (1979) adjustment instead of weighting coefficients equally (which is done in OLS), weights the coefficients depending on their accuracy (variance) when summing across cross-sectional regressions and could be considered as Weighted Least Squares (WLS) methodology. Rejection of

According to Miller and Scholes (1982) and Berk (1995), possible addition variables (size, book-to-market, turnover, etc.) are related to stock price; therefore, cross-sectional regressions might capture some mis-measurment of market beta rather than the independent pricing factor. Since the PIN variable is solely related with the number of transactions, it avoids possible critique.

4 Data

4.1 Data for PIN Estimation

EHO (2002) classified all trades as buy trades or sell trades based on Lee-Ready (Lee & Ready, 1991) algorithm which states that any trade that took place above (below) bid-ask midpoint can be regarded as a buy (sell) trade because trades initiated by buyers (sellers) usually are executed at or near the ask (bid). This method might induce a bias because of a trade misclassification; therefore, in this research it is not applied. Since the Baltic stock exchanges are order-driven with buy and sell orders submitted and auctioned automatically as they arrive, it provides 100% correct trade direction. The data set of daily number of total buy trades and sell trades consists of 21,322 data points for the VSE (sample period from July 2005 to October

2008, T=13), 10,694 data points for the RSE (sample period from October 2004 to October 2008, T=16), and 13,480 data points for the TSE (sample period from October 2004 to October 2008, T=16). The data set was obtained from NASDAQ OMX. However, the drawback of this data set is that it does not include a part of trades that were matched during the opening or closing auctions as there is no secondary party that can be identified. Due to this reason 10.22% of trades in the VSE, 8.35% of trades in the RSE, and 13.80% of trades in the TSE did not have classification of a buy or sell trade. These trades are omitted from the research. The data set covers relative short time period (because more data is not available on the Baltic markets). However, since all asset pricing tests are asymptotic (assumes infinity), this data set provides as high validity results as large data would present. The much more important thing is to ensure that there are not structural breaks and extreme observations in the sample. Thus, the data sample is studied with a great attention (see sections 5.1 – 5.2).

The data set includes 32 listed companies in the VSE, 13 listed companies in the RSE, and 14 listed companies in the TSE. Other companies did not meet minimum liquidity requirement set in this study (at least 20 trading days per quarter). The list of companies and their descriptive trading statistics can be seen in the Table 1 (Appendix C).

4.2 Data for interrelationship analysis

Asset pricing tests required some additional data on returns and firm characteristics for companies in the Baltic equity market. Quarterly firm returns, controlled by stock splits and other corporate events, were obtained from NASDAQ OMX. They were also adjusted for dividends and bonus issues. In order to estimate equity betas, one month money market interest rates, and returns on Vilnius, Riga and Tallinn stock markets indices were chosen. Short term interest rates were applied because of two reasons: 1) they do not convey the interest rate risk that is incorporated in long-term interest rates; and 2) there is no data on long term interest rates in Estonia since Estonian government runs budget surplus. In order to capture the *SIZE* variable, firm's data on capitalization in EUR were collected from NASDAQ OMX. Book-to-market ratios were obtained from Reuter's database.

5 Results

The following section overviews main results of the study. In order to control that data set does not include structural breaks, firstly, the parameters that were discovered by the maximum likelihood procedure are described. Secondly, asset pricing variables and their descriptive statistics are overviewed. Lastly, the interrelationship between variables is investigated applying portfolios sorted by PIN, the correlation matrix and regressions.

5.1 PIN Estimation

5.1.1 Boundary Solutions

Yan and Zhang (2006) report that the estimation of maximum likelihood function parameters might involve a bias related to corner solutions. When authors refer to corner solutions they mean that there might be too many cases when function parameters α (the probability that information event happens) and δ (the probability that information conveys a bad signal to the market) are equal to 1 or 0 which imply higher dispersion and bias of PIN variable. Therefore, this subsection investigates corner solutions and verifies if they provide an economic rationale or should be omitted.

For parameter α, 50 procedure solutions fell on the boundary of the parameter space, out of which 31 were for VSE, 11 for RSE, and 8 for TSE. The majority of these cases (42 out of 50) indicate that for some stocks on certain quarters, information events were happening each day. On the contrary, 8 cases demonstrate that for some stocks private information events did not occur during particular quarters at all. Yan and Zhang (2006) question whether these corner solutions induce a bias in PIN estimation, since public information disclosures should produce some private information ($\alpha \neq 0$) and it is a rare case that information events occur each day in developed markets ($\alpha \neq 1$). However, looking closer at the sample of this study, the rationale for boundary solutions can be supported. For example, the procedure generated corner solutions ($\alpha = 0$) for Mazeikiu Nafta (ticker: MNF1L) on 2007Q1 and 2007Q2. This comes in line with the fact that on 15th of February 2007, the squeeze-out of company's shares was announced. Thus, there was no private information events happening till the end of all share repurchase. Since Baltic equity markets are considered as maturing, the argument for full information events also loses a power. For parameter δ, there were 75 corner solutions that arise because of sustained imbalance of trading, out of which 36 were for VSE, 14 for RSE, and 25 for TSE. The majority (56 out of 75) of them were related with some stocks that on certain quarters conveyed only bad information to the market. All in all, relatively low number of boundary solutions for both α (6.7% of all cases) and δ (9.9% of all cases) indicate that maximum likelihood procedure provides economically reasonable results.

5.1.2 Distribution of Parameters

Table 2 (Appendix C) contains the summary statistics of maximum likelihood procedure parameter estimates and the PIN. In total the maximum likelihood procedure generated 759 (392 for VSE, 186 for RSE, and 181 for TSE) PIN observations on quarterly basis. The pooled panel PIN is 0.339, which is comparably higher than EHO (2002) PIN of 0.191 on NYSE, Lu and Wong (2007) PIN of 0.201 on Taiwan Stock Exchange, and similar to Heidle and Huang (2002) PIN of 0.333 on NASDAQ stocks that move to NYSE. However, these results cannot be directly compared between different markets since the time period and sample collection methods do not match. Nevertheless, PIN estimates for Baltic equity markets display the same sort of distribution as in previous studies (see Figure 2 in Appendix D). In terms of Baltic countries, Estonian stock market has the lowest probability of information-based trading (PIN value of

0.318), followed by Lithuanian (PIN value of 0.338), and further by Latvian (PIN value of 0.364) equity markets. Even though it is usually assumed that all three Baltic stock markets are mutually the same, these results indicate that there are significant differences among markets, at least when it is referred to information-based trading. There might be several reasons for this phenomenon. First, it might be a case that Lithuanian and Latvian stocks are traded relatively more by informed traders or, in other words, there are relatively less liquidity traders compared to Estonian market. Second explanation could be that information events happen more frequently in RSE than in other two Baltic markets.

Overall, significantly higher PIN estimates demonstrate that Baltic stock markets, similarly to the case of Prague Stock Exchange (Hanousek and Podpiera, 2002), could be associated with higher levels of insider trading that are common in emergent equity markets. Higher probability of information-based scores should be particular important to policy makers and market regulators since they are main bodies that should ensure high transparency and fair trade in a market. In order to reduce the information asymmetry and improve PIN scores regulatory bodies should set higher disclosure requirements and promote voluntary disclosures. According to Brown et al. (2004) this would directly lower the fraction of private information to total information available in the market and indirectly reduce incentives to seek for private information.

Figure 3 in Appendix D depicts the development of 25th, 75th percentiles and the mean of PIN across quarter. As it can be seen, the probability of information-based trading is rather stable across time. Only 25th and 75th percentiles depict larger variations that are due to boundary solutions in α estimate. PIN is also found to be negatively correlated cross-sectionally with trading turnover for almost all quarters (see correlation matrix of the PIN and turnover in Table 5 in Appendix H). The fact, that the PIN almost does not vary across time but at the same time is negatively correlated with turnover across stocks (which is consistent with Easley, Kiefer, O'Hara and Paperman (1996)), implies that stocks that are traded more frequently are more likely to have less adverse selection problems due to informed trading. In other words, this means that information asymmetry between informed and uninformed investors decreases if a stock is traded more often. This is because of the fact that with each additional transaction and price movement investors are able to extract more and more information about a stock and in such way update their beliefs about a true value of it. It is expected that turnover and the PIN has a circular cause-and-effect. This means that higher turnover of a stock leads to smaller information asymmetry which further imply higher turnover. Thus, continuous efforts to reduce the probability of information-based trading (by policy makers or individual companies) in long term should result in sustainable growth of the market.

The parameter α (the probability of information event happening) does not vary much across time with an overall pooled mean of 0.360 and standard deviation of 0.224 (see Figure 5 in Appendix E). Only second quarter scores are on average a bit higher than other quarters. It is line with economic rationale since on second quarter listed companies in the Baltics usually announce annual results that cause more discussions and information signals than announcement of interim results. The mean of α for pooled data set is significantly higher than the mean of 0.283 in EHO (2002) indicating that information events occur more often in the Baltic stock markets than in NYSE. One of reasons for this difference could be that in the Baltics there is less stocks traded; thus, information events are more concentrated. Even though the distribution of α for the Baltic equity markets is more dispersed, it is skewed to the right as for other markets (see Figure 4 in Appendix E). On a country level, RSE appears to have the highest probability of information event happening.

On the contrary from α, the distribution of pooled sample of δ parameter has a negative skew (see Figure 6 in Appendix F) indicating that it is more likely that any information event in Baltic stock markets, on average, had a negative signal to the market. It is believed that this characteristic is only sample-dependent and cannot be accepted as a rule: there is no reason why firms listed on the Baltic stock exchanges should produce more negative news to the market. Parameter δ averages across the time are illustrated in Figure7 in Appendix F. Estimates of parameters ε_b , ε_s and μ also provide an economic intuition. Parameter estimated values are highly positively correlated with turnover which corresponds to the idea that higher arrival rates of traders generate larger turnover. The correlation coefficients of ε_b , ε_s, μ and turnover are summarized in Table 6 in Appendix H.

Since the probability of information-based trading is not observable, one might cast a doubt that the PIN is not measuring what it is intended to measure. However, parameter estimates, that are consistent with the economic rationale, imply that the PIN calculation is correct and the PIN is an appropriate measure of information risk.

5.2 Summary Statistics of Asset Pricing Variables

The collected data set covers a broad range of companies, from the smallest, with market capitalization of EUR 0.8m, to the largest, with market capitalization of EUR 2.6bn, while an average company has a market capitalization of EUR 191.5m. VSE and TSE have significantly larger companies (the average capitalization of EUR 212.9m and EUR 212.6m, respectively) than RSE (the average capitalization of EUR 125.7m). The book-to-market ratio varies substantially as well: from 0.1 to 9.2 with an average value of 0.95. RSE includes more “value stocks” (average book-to-market value of 1.34) while TSE has more “growth stocks” (average book-to-market value of 0.51). Over the sample period, equally-weighted portfolio of TSE stocks on average yielded highest excess returns of -1.25% per month compared to -4.32% and -4.77%

for RSE and VSE, respectively. The average stock in the pooled sample has a beta estimate of 0.84. Asset pricing variables statistics are summarized in Table 3 Appendix G.

5.3 Interrelationship of Variables

In this section the interrelationship of variables is presented. Firstly, company characteristics are viewed for two portfolios. Secondly, the correlation matrix of all variables is investigated. Lastly, since correlation of variables does not necessarily imply causation between variables, regression analysis is performed.

5.3.1 Portfolios PIN Low and PIN High

Since the interest of this paper lies within the variable PIN, this section overviews if any conclusions about PIN could be made considering specific company characteristics. For this reason all stocks of the sample are sorted into two portfolios: PIN Low, which includes all stock-quarters that have PIN below the median value, and PIN High, which includes all stock-quarters that have PIN above the median value. Table 4 in Appendix G reports summary statistics for asset pricing variables sorted by these two portfolios. It is found that portfolios have different average excess returns; however, they are not significantly different from each other. This finding is a first indicator that PIN might not be priced in the Baltic equity markets (rejection of H4) since higher PIN value does not correspond to significantly higher excess return. The analysis shows that stocks-quarters of portfolio PIN Low have significantly (at 5% significance level) higher betas (average beta value of 0.89) than those of portfolio PIN High (average beta value of 0.80). These results indicate that companies with greater private information have lower systematic volatility. The comparison also reveals that large stock-quarters have significantly (at 0.1% significance level) lower PIN estimates which means that it is more difficult for informed traders to extract private information from large companies than from small companies. Since large companies usually have better disclosure and, in general, the interest of society is usually more active towards big companies, it becomes difficult for informed traders to obtain advantageous private information. This finding is also consistent with Easley, Kiefer, O'Hara and Paperman (1996). In addition, companies of portfolio PIN Low have significantly lower book-to-market ratio than companies in PIN High portfolio.

Portfolio analysis reveals that already from only few company characteristics it is possible to draw general conclusions about the expected level of information asymmetry of a stock. Unfortunately, no relationship is found between PIN and excess return since the difference of excess returns of portfolios is found insignificant. Still, summary statistics depict that higher PIN scores are present for small and growth stocks. In addition, Companies that have higher systematic risk tend to have the lower probability of information-based trading. Data on debt-equity ratio could reveal if such conclusions could be made about business risk as well. However, it is out of scope of this research and could be a topic in further research.

5.3.2 Correlations

Table 7 in Appendix H summarizes cross-sectional correlations between asset pricing variables. Correlation matrix provides another argument for interrelationship between PIN and other variables. In addition, it also presents how asset pricing variables are related between each other. The correlation coefficients support the previous findings of portfolio analysis: PIN is negatively correlated with the size of a company (as expressed by market capitalization), positively related with the book value over market value of a company, and has no interrelation with excess returns. This provides another indication that PIN might not be priced in the Baltic stock markets. The previous conclusions about company beta and PIN lose a power since the correlation coefficient between those two variables is not significantly different from zero. The correlation analysis also reveals that the interrelationship between excess returns and *SIZE* and between excess returns and *BM* are insignificant. In other words, correlation matrix suggests that large and small, or value and growth stocks should deliver the same excess return. Moreover, beta estimates are found to be significantly related with excess returns. However, surprisingly, the coefficient has a negative sign of -0.187 indicating that higher systematic risk is not rewarded with greater returns. Even though this finding contradicts the economic rationale, this is in line with previous research on the Baltics. Pajuste (2002) shows that the risk-return trade-off is inverse in Central and Eastern Europe and critiques common emerging market characteristic (high risk is associated with higher returns) because of survivorship bias. Inverse risk-return trade-off anomaly was observed in developed equity markets as well (see Fama & French (1992), Chalmers & Kadlec (1998), or Datar et al. (1998)). The correlation analysis also reveals that stocks with higher systematic volatility tend to be larger (correlation coefficient for β and *SIZE* of 0.161) and have higher book-to-market ratio (correlation coefficient for β and *BM* of 0.154).

5.3.3 Regression Analysis

For regression analysis the time-series averages of coefficients of cross-sectional asset pricing tests are applied following the standard Fama-MacBeth (1973) methodology with Litzenberger and Ramaswamy (1979) Weighted Least-Square adjustment. Regression results are summarized in Table 8 in Appendix J. First the regression analysis is performed for each Baltic stock market separately, and afterwards for pooled data set.

Regression analysis for VSE. The results of the regression on VSE are summarized in Panel A, Table 8 (Appendix J). Regression results reveal that the probability of information-based trading is a significant determinant (t-value of 1.43) of stock returns in Lithuanian equity market: 10% increase in PIN score results to positive increase of 3.8% in excess returns. This indicates that investors in Lithuania, that do not have the advantage of private information, require higher return for stocks that convey such information. In other words, information asymmetry influences stock price evolution and creates additional risk of holding it. However, PIN significance cannot be rejected only at 10% significance level. Therefore, one should be cautious about power of this result.

Regression analysis also reveals that there is significant relationship between excess returns and beta (t-value of 3.36) in VSE. However, as in correlation matrix, the coefficient has a negative slope: a stock with beta coefficient that is 0.5 higher delivers 4.66% lower expected return. Thus, there is an inverse risk return trade-off. Moreover, additional variables suggested by Fama and French (1992), *SIZE* (t-value of 3.50) and *BM* (t-value of 2.91), also found to be significant determinants of stock returns in VSE. One percent increase in stock capitalization leads to 2.33% higher excess return, and one percent increase in book-to-value ratio results in 1.02% increase in excess return.

Overall, the regression for VSE indicates that the probability of information-based trading is weakly significant determinant of stock returns in Lithuanian equity market; therefore, the first hypothesis (H1) cannot be rejected. Additionally, analysis shows that investors accept lower return for higher systematic risk which at first glance does not follow the economic rationale. Furthermore, it is found that large and “value” stocks tend to deliver higher returns than small or “growth” stocks.

Regression analysis for RSE. Panel B in Table 8 (Appendix J) summarizes regression results for RSE. Regression analysis indicates that, contrary to VSE, stock returns in Latvia are not driven by the probability of information-based trading. Even though the coefficient has a sign as expected, it is not different from zero (t-value of 0.66). Thus, information asymmetry does not influence stock returns in RSE. Similarly to VSE, other factors, beta, *SIZE* and *BM* are found to be highly significant; however, the extent they affect excess returns is different from VSE. In comparing two stocks, the stock with higher beta coefficient of 0.5 would deliver 2.10% lower expected excess return in RSE. Moreover, 1% percent increase in market capitalization on average corresponds to 1.47% excess returns. Furthermore, 1% increase in a book-to-value ratio leads to 4.57% rise in expected excess returns.

Overall, regression results on RSE reveal that uninformed traders do not require higher returns for holding stocks that convey more private information. Therefore, the probability of information-based trading is not a determinant of stocks returns in RSE and second hypothesis (H2) of this research is rejected. Moreover, investors in Latvian equity market do not require higher return for higher systematic risk; they even accept lower return for higher market risk. It is also found that large and "value" stocks outperform small and "growth" stocks, respectively.

Regression analysis for TSE. The output of the regression for TSE is summarized in Panel C, Table 8 (Appendix J). Regression results illustrate that similarly to the case of RSE, excess stock returns in Tallinn equity market are not affected by the probability of information-based trading. The PIN coefficient has a required slope sign; however, its influence on expected excess returns is insignificant (t-value of 0.34). This indicates that information asymmetry does not affect stock prices in Tallinn market. Similarly to VSE and RSE, investors in Estonian market accept higher risk and lower return: 0.5 unit increase in beta coefficient on average results in 5.41% reduction in excess returns. In addition, "value" stocks tend to outperform "growth" stocks: 1% increase in book-to-market ratio corresponds to 1.47% increase in expected excess returns. On the contrary to VSE and RSE, *SIZE* factor is found to be insignificant (t-value of 0.44).

Overall, regression analysis for TSE reveals that investors do not require additional return for holding stocks that convey more private information in Estonian market. Therefore, the probability of information-based trading cannot be considered as a pricing factor of stock returns in TSE and third hypothesis (H3) of this research is rejected. Similarly to other Baltic markets, investors in TSE accept lower returns with higher systematic volatility and consider “value” stocks as more expensive.

Regression analysis for pooled data set. In order to fade out a doubt that regression results might be biased because of the amount of data for each of the markets, this research reports regression results for pooled data set, that are summarized in Panel D, Table 8 (Appendix J). Results indicate that the probability of information-based trading does not affect expected stock returns in Baltic equity markets. The coefficient for PIN is found to be insignificant from zero (t-value of 0.39). Not surprisingly, the systematic risk is found to be wrongly priced in pooled data as well. The increase of 1 unit in beta factors leads to 6.30% reduction in expected stock returns. Moreover, other factors suggested by Fama and French (1992), *SIZE* and *BM*, have a significant impact on average stock excess returns. One percent increase in the market capitalization of a stock results in 1.50% higher returns, while one percent rise in book-to-market ratio corresponds to 0.32% increase in excess returns.

In general, the analysis for pooled data set reveals that information asymmetry is not priced in overall Baltic equity market. Therefore, the fourth hypothesis (H4) of this research is rejected. It also proves that the return risk trade-off anomaly which is observed in all three separate markets is not accident and is present in subsumed analysis as well. Moreover, it is found that large and value stocks deliver higher expected excess returns. Thus, the regression results for pooled data present almost the same significance of coefficients as in each separate market. Therefore, it provides another robustness check and validity of regression analysis.

5.4 Discussion of Results

The average PIN score for the pooled data set is 0.339. This figure is significantly higher than for developed stock markets and similar as in Prague Stock Exchange. This signifies that information asymmetry levels are much higher in emerging equity markets. Partly, this also indicates that insider trading is more common in the Baltic equity markets. In addition, significant differences of the PIN are observed in separate Baltic markets. There might be several reasons for that. First, it could be a case that Lee-Ready (1991) algorithm used in previous research causes downward bias in PIN in as noted by Boehmer et al. (2007). Since this paper does not apply this algorithm, naturally PIN scores could look relatively higher. Second the upward bias in PIN could be caused by large number of boundary solutions when certain stocks signal news very frequently ($\alpha = 1$); however, the investigation of corner solutions indicate that all of them are in line with economic rationale and cannot be omitted. Third, information events in one equity market happen more frequently than in others, which in turn depend on how much companies are transparent. Since information events happen most frequently in Latvia it is not a surprise that its equity market has a highest PIN score. Fourth, the higher share of informed traders also leads to higher information asymmetry. From this perspective, in both VSE and TSE informed traders have lower fraction. Finally, company characteristics also matter. According to Brown et al. (2004) higher voluntary disclose should reduce information asymmetry. Even though there is no research on disclosure levels in Baltic Stock Exchanges, it would not be a surprise that Estonian and Lithuanian companies disclose more. In addition, correlation results reveal that large and "growth" companies in the Baltics have the lower probability of information-based trading. This also might be due to the fact that it is more difficult for informed traders to extract private information from such companies since they tend to disclose more. Furthermore, this analysis reveals that stocks with higher turnover on average have lower asymmetric information levels. Overall, it is believed that all these factors determine that Estonian market has the lowest probability of information-based trading followed by Lithuanian and further by Latvian markets.

After investigation of interrelationship of variables it is found that the PIN is only weakly priced in Vilnius stock exchange while in the others and overall Baltic equity market it is found to have no effect on expected stock excess returns. Thus, H2, H3, H4 are strongly rejected, while H1 cannot be rejected at 10% significance level. There might be possible several market specific reasons why the PIN is priced in VSE but not in RSE and TSE. Firstly, it could be because of some different trading rules among markets. However, NASDAQ OMX applies the same requirements for traders that want to trade in VSE, RSE or TSE. Thus, this argument is not relevant in this case. Secondly, it could be because of some demand factors; for instance, the activity of member firms in NASDAQ OMX. Taking a look to the period of 2005Q1 – 2008Q4, the NASDAQ OMX trading data shows that in all three Baltic exchanges the majority of trading activity was generated by several big players. Estonian market has a greatest concentration – 93% of trading turnover was generated by 3 companies (that have contributed to 5% or more of all trading activity in TSE); followed by Latvian market – 87% of trading turnover by 5 companies (that have contributed to 5% or more of all trading activity in RSE); and Lithuanian market – 71% by 4 companies (that have contributed to 5% or more of all trading activity in VSE). Thus, Lithuanian market has the highest fraction of small traders. Since usually big players have better access to information and in some cases private information (road shows, analyst coverage, etc.) it is expected that exactly these small traders demand a compensation for holding stocks that convey private information. Since the share of small traders is the biggest in VSE it might be the case why the PIN is weakly priced in Lithuanian equity market but not in RSE or TSE. It is expected that the similar conclusions could be drawn by looking at mutual and pension funds and their assets invested relatively to all market capitalization. Unfortunately, such data is not available. Another demand factor could be the perception of existing risk in different markets. As Hughes et al. (2005) note investors might subsume the information risk into other risk components. Thus, in the case of this research, investors investing in RSE and TSE might accumulate information risk into other variables. Second group of factors, that might be important why the PIN is priced in Lithuania but not in Latvia and Estonia, is supply factors. For example, the level of disclosure. If companies in Lithuania disclose relatively less it further leads to higher asymmetry level. Thus, uninformed traders require additional compensation for holding stocks that have higher the probability of information-based trading. Another factor could be firm characteristics. Perhaps, investors require compensation only for stocks that operate in particular industries or are of certain size. Thus, additional company specific characteristics could improve explanation of stock returns and check if the PIN remains or appears to be significant.

Even though the primary interest of this research lies within the PIN, regression analysis revealed some interesting facts that are worth discussing. First of all, it is found that in all three markets investors do not require higher returns for higher systematic volatility of stocks. On the opposite, they even accept lower expected returns for higher systematic risk. There might be couple of reasons for that. Firstly, the inverse risk return trade-off could be time specific, in other words it could be a case that for 2005Q1 - 2008Q4 there were some anomalies in stock pricing that could not be explained by Capital Asset Pricing Model. For example, Fama and French (1992) find that certain periods experienced negative return and market relationship in the US stock markets. Secondly, the inverse risk return trade-off might be present if risk free returns are higher than expected stock returns. Since average excess stock returns in VSE, RSE and TSE were negative (see summary statistics in Table 3, Appendix G), this is exactly the reason why slope of Security Market Line (SML) is negative. After adjusting for negative expectations, Pettengill et al. (1995) find that return and beta relationship is positive for the sample where without the adjustment the relationship was negative. Thus, the further research should apply Pettengill et al. (1995) adjustment and see if it improves results.

The regression analysis also prove that stock risks are multidimensional: stock returns in the Baltics are found to be driven by size and book-to-market (except for TSE) factors. Even though it is a common to find these two variables to be significant determinants of stock prices the only argument for their significance is rational pricing story. Fama and French (1992) argue that *BM* is priced because stocks that market perceive them as having poor prospects should have higher returns because they are penalized with higher cost of capital. Another reason could be that size and book-to-market follows the evolution of profitability and, thus, they are priced as well (Fama and French (1995)). However, the finding, as it is in this research, is purely empirical and lacks economic rationale.

6 Conclusions

Following the increasing recognition and application of market microstructure models, this research provides further evidence on market mechanism and its explanatory power for asset returns in the Baltic equity markets. The paper applies sequential trade model developed by Easley, Hvidkjaer & O'Hara (EHO) (2002). The main idea behind the model is that uninformed investors because of asymmetric information are always put into disadvantageous situation because of which they require higher expected returns. The model provides parameter estimates out of which the PIN, a measure of the probability of information-based trading is constructed. Its significance for stock returns is tested using Fama-MacBeth (1973) two step methodology under Litzenberger and Ramaswamy (1979) WLS adjustment. Additional factors, proposed by Fama and French (1992), size and book-to-market ratio, are included into regressions which are run for each Baltic market separately and for pooled data set.

The parameter estimates of EHO (2002) model for the Baltic stock markets are found to be economically reasonable; therefore, even though the PIN is not directly observable in the market, it provides a proper measure for information risk. The results indicate that the average fraction of trades that are information-based is 0.339 in the pooled data set. Compared to developed markets, the Baltic stock exchanges have much higher levels of information asymmetry. Higher levels of insider trading and lack of voluntary disclosure are possible reasons for that. Out of three Baltic stock exchanges Tallinn market is found to have the lowest PIN of 0.318, followed by Vilnius (PIN of 0.338) and Riga (PIN of 0.364). The differences of PIN scores among the Baltic markets are mainly influenced by relative number of informed traders and frequency of information events which are further predetermined by the profile of traders and company characteristics (transparency, size, etc.). These implications of findings should be particular important to policy makers who should put additional effort to ensure equal access to information for all market participants.

The interrelationship of excess returns and the PIN is investigated by analyses of two portfolios sorted by the PIN, the correlation matrix and regressions. Results are highly robust since all methods provide similar findings. The PIN is found to be weakly priced (at 10% significance level) in VSE: 10% increase in PIN score corresponds to 3.8% increase in expected excess return. In RSE, TSE and overall Baltic market data the PIN has no effect on excess stock returns. These findings imply that hypotheses H2, H3, H4 are rejected while H1 is not. Even though it is generally assumed that all three Baltic markets are mutually the same, results of this paper indicate that it is not true, at least when it is referred to information risk. Several demand and supply factors could be possible reasons for differences identified. First, the concentration of trading houses is much more predominant in TSE and RSE, or in other words, there are more small traders in VSE. It is believed that these small traders have worse or no access to private information because of which they demand higher returns. Another reason could be that investors in VSE consider information risk as a separate type of risk while investors in RSE and TSE include it into other factors. From the supply point of view, it is expected that certain company features (corporate governance, ownership, industry, capital structure, etc.) might influence if the PIN is priced.

The analysis also revealed some interesting facts about other determinants of stock returns. In all three Baltic stock exchanges there is an inverse relationship of market risk and return: investors accept lower returns for higher risks. One reason for that could be that it is only a time specific anomaly. Another, much more feasible, is that risk free rates were higher than stock returns, thus SML has a negative slope. Variables, suggested by Fama and French (1992) also are found to be significant. Large and "value" companies tend to outperform small and "growth" companies. Even though these variables are identified in other markets as well, the academic world still lacks rationale for them. Thus, this analysis also provides just empirical evidence on size and book-to-market effects.

The further research on topic should be directed to provide new insights and strengthen the power of this paper results. New perspective towards the PIN and Baltic stock markets could be gained by dividing the PIN into the part that is related to illiquidity and into the part that is related to information asymmetry as described in Duarte and Young (2009). In addition, deeper investigation what predetermines PIN levels in the Baltics should be made. Perhaps, for the first steps, one should follow Brown et al. (2004). This would directly identify fields that are necessary to work on to ensure equal access to information. From broader perspective, new insights on topic could be gained if the PIN model is applied to other Central and Eastern Europe. The same selection and model specifications would allow comparing different levels of information asymmetry among other emerging markets. Considering the probability of information-based trading and asset returns, further research on the Baltics should be linked with additional controlling variables in regression analysis, different model and time specifications. This would ensure high robustness of empirical results.

References

Agarwal, P., & O'Hara, M. (2006). Information Risk and Capital Structure. *Working Paper*.

Aktas, N., Bodt, E., Declerck, F., & Oppens, H. (2007). The PIN anomaly around M&A announcements. *Journal of Financial Markets,* 10(2), 169-191.

Amihud, Y. (2000). Illiquidity and Stock Returns: Cross-sectional and Time Series Effects. *Working Paper*.

Amihud, Y., & Mendelson, H. (1986). Asset Pricing and the Bid-Ask Spread. *Journal of Financial Economics,* (17), 223-249.

Amihud, Y., & Mendelson, H., (1989). The Effects of Beta, Bid-Ask Spread, Residual Risk, and Size on Stock Returns. *Journal of Finance,* 44(2), 479-486.

Amihud, Y., Mendelson, H., & Lauterbach, B. (1997). Market microstructure and securities values: Evidence from the Tel Aviv Stock Exchange. *Journal of Financial Economics,* 45(3), 365-390.

Berglund, T., & Knif, J. (1999). Accounting for the accuracy of beta estimates in CAPM tests on assets with time-varying risks. *European Financial Management*, 5, 29-42.

Berk, J. B. (1995). A critique of size-related anomalies. *Review of Financial Studies,* 8(2), 275-286.

Berk, J. D. (2000). Sorting Out Sorts. *Journal of Finance,* 55, 407-427.

Bharath, S., Pasquariello, P., & Wu, G. (2008). Does Asymmetric information Drive Capital Structure Decisions? *Review of Financial Studies*.

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

Brannas, Kurt, & Soultanaeva (2006). Influence of News in Moscow and New York on Returns and Risks on Baltic State Stock Indices. *Umea Economic Studies* (696).

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

Brown, S., Hillegeist, S. A., & Lo, K. (2004). Conference Calls and Information Asymmetry. *Journal of Accounting and Economics,* 37(3), 343-366.

Brunnermeier, M. (2001). *Asset Princing under Assymetric Information: Bubbles, Crashes, Technical Analysis, and Herding*. New York: Oxford University Press Inc.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The Econometricsof Financial Markets*, Princeton: Princeton University Press.

Chalmers, J.M., & Kadlec, G. B. (1998). An empirical examination of the amortized spread. *Journal of Financial Economics,* (48), 159-188.

Conrad, J., Cooper, M., & Kaul, G. (2003). Value versus Glamour. *Journal of Finance,* 58, 1969-1995.

Datar, V. T., Narayan, Y. N., & Radcliffe (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets*, (1), 203 – 219.

Devyzis, L., & Jankauskas, G. (2004). Explaining the Cost of Equity in Central and Eastern Europe. *SSE Riga working papers,* 13(68).

Duarte, J., Han, X., Harford, J., & Young, L. (2006). Information asymmetry, information dissemination and the effect of regulation FD on the cost of capital. *Journal of Financial Economics,* 87(1), 24-44.

Duarte, J., & Young, L. (2009). Why is PIN priced? *Journal of Financial Economics,* 91(2), 119-138.

Easley, D., Hvidkjaer, S., & O'Hara, M. (2002). Is Information Risk a Determinant of Asset Returns? *Journal of Finance,* 57, 2185-2221.

Easley, D., Hvidkjaer, S., & O'Hara, M. (2005a). Factoring information into returns. *Working paper.*

Easley, D., Hvidkjaer, S., & O'Hara, M. (2005b). Information and the Cost of Capital. *Journal of Finance,* 59(4), 1553-1583.

Easley, D., Kiefer, N. M., & O'Hara, M. (1996), Cream-skimming or profit sharing? The curious role of purchased order flow, *Journal of Finance,* (51), 811-833.

Easley, D., Kiefer, N.M., & O'Hara, M. (1997a). The information content of the trading process, *Journal of Finance,* (4), 159-186.

Easley, D., Kiefer, N.M., & O'Hara, M. (1997b). One day in the life of a very common stock. *Review of Financial Studies,* 10(3), 805-835.

Easley D., Kiefer, N. M., O'Hara, M., & Paperman, J.B. (1996). Liquidity, Information, and Less-Frequently Traded Stocks, *Journal of Finance,* 51(4), 1405-36.

Easley D., O'Hara M., & Paperman, J.B. (1998). Financial Analysts and Information-based Trade, *Journal of Financial Markets,* 1(2), 175-201.

Eleswarapu, V. R., & Reinganum, M. R. (1993). The seasonal behavior of liquidity premium in asset pricing. *Journal of Financial Economics,* (34), 373-386.

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance,* 47(2), 427-465.

Fama, E. F., & French, K. R. (1995). Size and Book-to-Market Factors in Earnings and Returns. *Journal of Finance*, (50), 131 – 155.

Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy,* 81, 607-636.

Fuller, K., Van Ness, B., & Van Ness, R. (2007). Is Information Risk Priced for NASDAQ-listed Securities. *Working paper*.

Glosten, L. R., & Milgrom, P. R. (1985). Bid, Ask and Transaction Prices in Specialist Market with Heterogeneously Informed Traders. *Journal of Financial Economics*, (14), 71-100.

Hanousek, J., & Podpiera, R. (2002). Information-driven trading at the Prague Stock Exchange. *Economics of Transition,* 10(3), 747-759.

Heidle, H. G., & Huang, R. G. (2002). Information-Based Trading in Dealer and Auction Markets: An Analysis of Exchange Listings. *Journal of Financial and Quantitative Analysis,* (37), 391-424.

Hogfeldt, P., Kepitis, G., & Pajuste, A. (2000). Risk Factors and Predictability of Stock Returns in Central and Eastern Europe. *EMERGING MARKETS QUARTERLY*.

Hughes, J., Liu, J., & Liu, J. (2005). Information, Diversification, and Cost of Capital. *Working Paper*.

Kim, D. (1995). The Errors in the Variables Problem in the Cross-Section of Expected Stock Returns. *Journal of Finance,* 50, 1605-1634.

Kohn, M. (1994). *Financial Institutions and Markets.* Singapore: McGraw-Hill, Inc.

Kopriva, F. (2008). Source of information-driven trading on the Prague Stock Exchange. *CERGE-EI Working Paper*.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica,* (53), 1315 – 1335.

Lambert, R., Leuz, C., & Verrecchia, R. (2006). Accounting Information, Disclosure, and the Cost of Capital. *Working paper.*

Lee, C.M.C., & Ready, M.J. (1991). Inferring Trade Direction from Intraday Data. *Journal of Finance,* (46), 733-746.

Litzenberger, R., & Ramaswamy, K. (1979). The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of Financial Economics*, (7), 163 – 196.

Lu, Y.C., & Wong, W.K. (2007). Probability of Information-based Trading as a Pricing Factor in Taiwan Stock Market. *Working Paper*.

Miller, M. H., & Scholes, M. S. (1982). Dividends and Taxes: Some Empirical Evidence. *Journal of Political Economy,* 90(6), 1118-1141.

Mohanram, P.S., & Rajgopal, S. (2007). Is Information Risk (PIN) Priced? *Working paper*.

Nemecek, L. (1997). Liquidity and information based trading on the order driven capital market: the case of the Prague Stock Exchange. *CERGE-EI Working Paper.*

Pajuste, A. (2002). Corporate Governance and Stock Market Performance in Central and Eastern Europe. *Working paper*.

Vega, C. (2006). Stock price reaction to public and private information. *Journal of Finance, Economics* 82, 103-133.

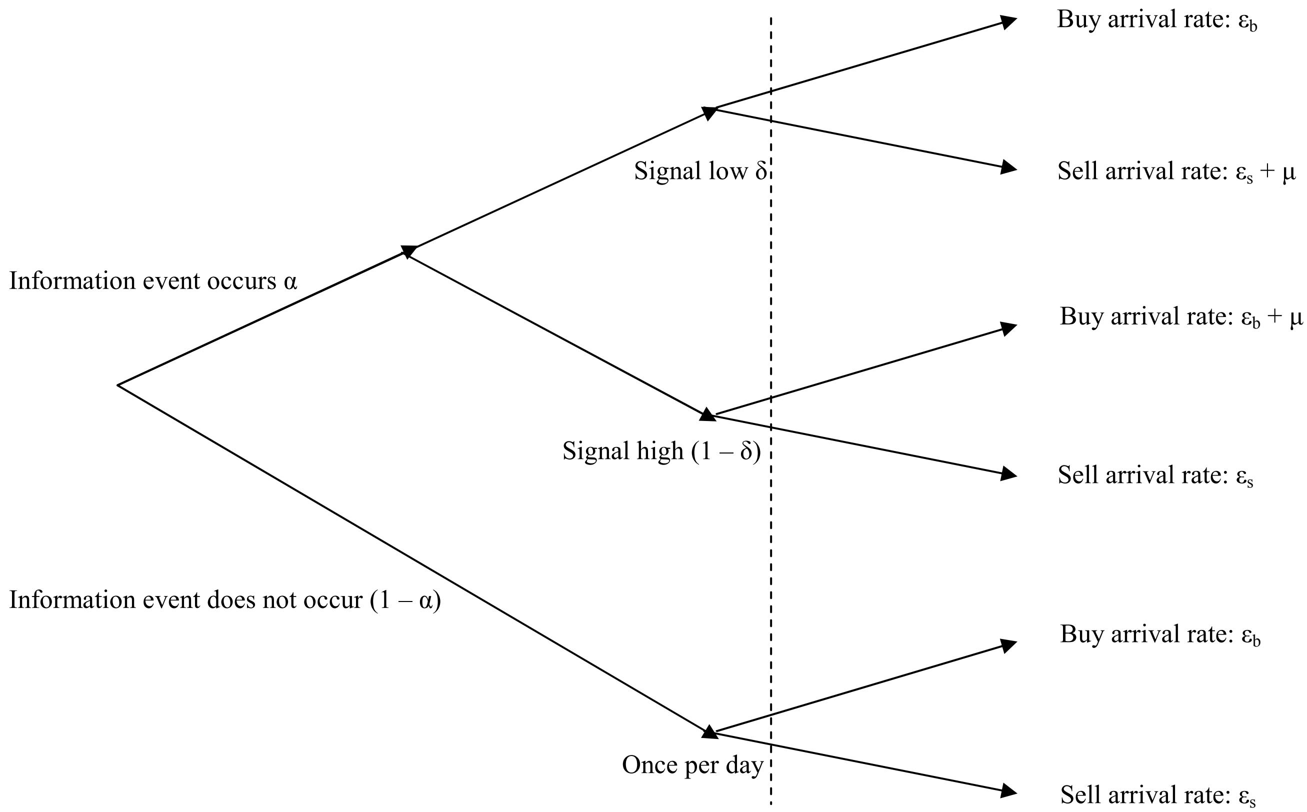
Wong, W.K., Tan, D., & Tian, Y. (2008). Informed Trading and liquidity in the Shanghai Stock Exchange. *International Review of Financial Analysis.*

Yan, Y. & Zhang, S. (2006). An Improved Estimation Method and Empirical Properties of the Probability of Informed Trading. *Working paper*.

Appendix A

Figure 1. *Tree diagram of the trading process*

α is the probability of information event occurring, δ is the probability that information conveys a bad signal, μ is the arrival rate of informed traders, ε_s and ε_b are the arrival rates of uninformed sell orders and buy orders, respectively.



Note: Easley, D., Hvidkjaer, S., & O'Hara, M. (2002). *Tree Diagram of the Trading Process*.

Appendix B

The SAS *NLP* procedure for numerical maximization of the factorized likelihood function

```
proc nlp data=stocks tech=newrap
maxstep=5 maxiter=100 maxfunc=250 ;
by ticker y quarter ;
max loglik ;
decvar ALPHA, DELTA, MU, eps_B, eps_S ;
M = min(B,S) + max(B,S) / 2.0 ;
X_B = eps_B / (mu + eps_B) ;
X_S = eps_S / (mu + eps_S) ;
negative_M = -1.0*M ;
negative_mu = -1.0*mu ;
one_minus_d = 1.0- delta ;
one_minus_alpha = 1.0- alpha ;
S_minus_M=S-M ;
B_minus_M=B-M ;
bounds 0.0 < eps_B eps_S mu, 0.0 <= alpha delta <= 1.0 ;
loglik =
-1.0*eps_B + M * log(X_B) + B*log(mu+eps_B)
-1.0*eps_S + M * log(X_S) + S*log(mu+eps_S)
+ log(
alpha*one_minus_d*X_S**S_minus_M*X_B**negative_M*exp(negative_mu)
+ alpha*delta*X_B**B_minus_M*X_S**negative_M*exp(negative_mu)
+ one_minus_alpha* X_B**B_minus_M* X_S**S_minus_M
) ;
run ;
```

Appendix C

Table 1

List of companies and their descriptive statistics

Stock Exchange	Ticker	Sample period from		Sample period to		Total number of buys	total number of sells	total number of trades	% of trades without identified direction
		year	quarter	year	quarter				
TSE	BLT1T	2004	4	2008	3	10521	11174	24396	11.07%
TSE	ETLAT	2004	4	2008	3	8361	8984	20015	13.34%
TSE	HAE1T	2004	4	2008	3	3235	3405	7495	11.41%
TSE	JRV1T	2004	4	2008	3	6206	6587	14674	12.82%
TSE	KLV1T	2004	4	2008	3	3646	3789	8352	10.98%
TSE	NRM1T	2004	4	2008	3	2650	2235	5592	12.64%
TSE	RLK1T	2004	4	2006	3	560	606	1265	7.83%
TSE	SFGAT	2005	1	2008	3	5026	5034	11462	12.23%
TSE	SKU1T	2004	4	2008	3	4100	3969	9159	11.90%
TSE	SMN1T	2005	3	2007	4	1801	1142	4076	27.80%
TSE	TAL1T	2005	4	2008	3	23792	21598	53155	14.61%
TSE	TKM1T	2004	4	2008	3	10840	11127	24843	11.58%
TSE	TPD1T	2004	4	2008	3	3004	3193	6887	10.02%
TSE	TVEAT	2005	3	2008	3	3581	3210	9758	30.41%
						87323	86053	201129	13.80%
VSE	ALT1L	2005	3	2008	3	2400	2102	5006	10.07%
VSE	APG1L	2005	3	2008	3	19340	17258	40441	9.50%
VSE	ATK1L	2005	3	2007	2	2264	2474	5585	15.17%
VSE	DKR1L	2005	3	2008	3	3270	3185	7108	9.19%
VSE	GRG1L	2005	3	2008	3	1613	1496	3457	10.07%
VSE	IVL1L	2005	3	2008	3	14105	12887	29872	9.64%
VSE	KBL1L	2005	3	2008	1	712	593	1520	14.14%
VSE	KNF1L	2005	3	2008	3	3958	4087	9100	11.59%
VSE	LBS1L	2005	3	2007	1	536	657	1372	13.05%
VSE	LDJ1L	2005	3	2008	3	3183	3190	7132	10.64%
VSE	LEL1L	2005	3	2008	3	794	1083	2145	12.49%
VSE	LEN1L	2005	3	2008	3	3775	4229	8989	10.96%
VSE	LFO1L	2005	3	2008	3	24300	23214	52087	8.78%
VSE	LJL1L	2005	3	2008	3	4248	3618	8674	9.32%
VSE	LLK1L	2005	3	2008	3	2337	2592	5432	9.26%
VSE	LNS1L	2005	3	2008	3	4499	4246	9653	9.41%
VSE	MNF1L	2005	3	2007	3	11762	12965	28075	11.93%
VSE	PTR1L	2005	3	2008	3	9695	9398	21445	10.97%
VSE	PZV1L	2005	3	2008	3	3326	3722	8223	14.29%
VSE	RST1L	2005	3	2008	3	7457	7821	16987	10.06%
VSE	RSU1L	2005	3	2008	3	3653	3016	7639	12.70%
VSE	SAB1L	2005	3	2008	3	18308	15420	37194	9.32%
VSE	SAN1L	2005	3	2008	3	5478	5718	12567	10.91%
VSE	SNG1L	2005	3	2008	3	2146	1668	4448	14.25%
VSE	SRS1L	2005	3	2008	3	12385	11118	26099	9.95%
VSE	STU1L	2005	3	2008	3	2734	2428	5711	9.61%
VSE	TEO1L	2005	3	2008	3	33417	31607	73945	12.06%
VSE	UKB1L	2005	3	2008	3	54465	49038	113290	8.64%
VSE	VBL1L	2005	3	2008	3	1188	780	2216	11.19%
VSE	VNG1L	2006	1	2007	4	962	989	2399	18.67%
VSE	VST1L	2005	3	2008	3	836	1162	2314	13.66%
VSE	ZMP1L	2005	3	2008	3	2381	2228	5194	11.26%
						261527	245989	565319	10.22%
RSE	BAL1R	2004	4	2008	3	3825	2515	6988	9.27%

RSE	DPK1R	2004	4	2008	3	5564	4542	10905	7.33%
RSE	GRD1R	2004	4	2008	3	4174	3063	8052	10.12%
RSE	GZE1R	2004	4	2008	3	2210	1962	4557	8.45%
RSE	LKB1R	2004	4	2008	3	4976	3389	9012	7.18%
RSE	LME1R	2004	4	2008	3	4261	3429	8403	8.49%
RSE	LSC1R	2004	4	2008	3	11472	10648	24158	8.44%
RSE	OLF1R	2004	4	2008	3	4224	4754	9725	7.68%
RSE	RKB1R	2004	4	2008	3	3352	3265	7149	7.44%
RSE	RTF1R	2004	4	2006	1	2346	2346	4993	6.03%
RSE	SAF1R	2004	4	2008	3	1865	1407	3684	11.18%
RSE	VNF1R	2004	4	2008	3	6036	3709	10726	9.15%
RSE	VSS1R	2004	4	2008	3	2198	1787	4377	8.96%
						56503	46816	112729	8.35%

Table 2

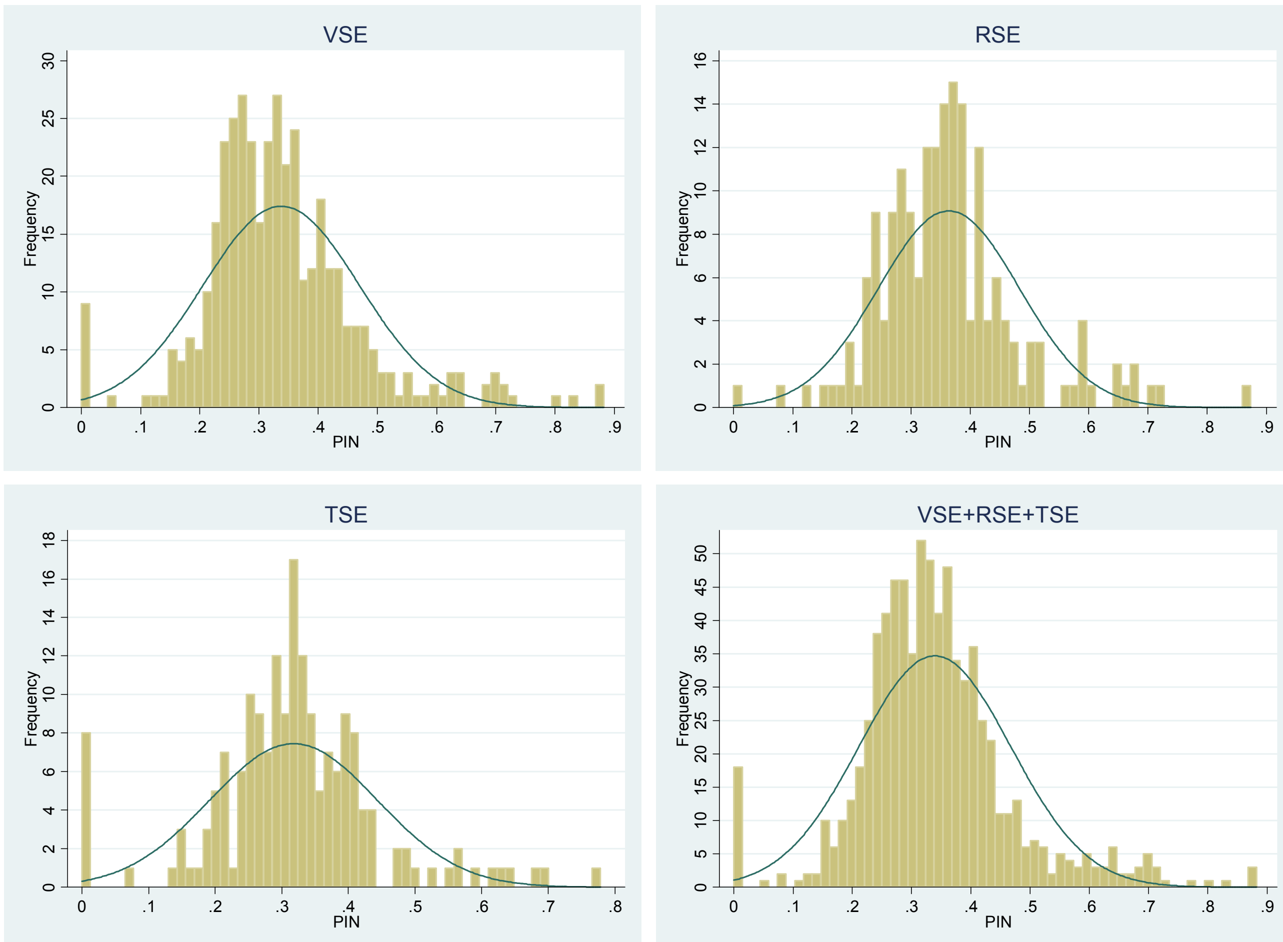
Maximum likelihood procedure parameter estimates

Pooled data (N=759)	**loglik**	**eps_b**	**mu**	**eps_s**	**delta**	**alpha**	**PIN**
Mean	183.226	6.392	16.490	5.383	0.563	0.360	0.339
St. Deviation	357.257	10.136	19.642	9.612	0.259	0.224	0.128
Median	60.565	3.108	10.049	2.420	0.574	0.317	0.328
Max	3035.487	89.458	171.509	105.673	1.000	1.000	0.883
Min	-1.917	0.000	0.000	0.000	0.000	0.000	0.000
VSE (N=392)	**loglik**	**eps_b**	**mu**	**eps_s**	**delta**	**alpha**	**PIN**
Mean	255.479	8.352	20.139	7.204	0.551	0.352	0.338
St. Deviation	449.849	12.870	23.495	12.109	0.250	0.227	0.132
Median	72.869	3.330	11.712	2.814	0.561	0.309	0.323
Max	3035.487	89.458	171.509	105.673	1.000	1.000	0.883
Min	-1.917	0.000	0.000	0.000	0.000	0.000	0.000
RSE (N=186)	**loglik**	**eps_b**	**mu**	**eps_s**	**delta**	**alpha**	**PIN**
Mean	77.624	3.598	10.344	2.408	0.581	0.382	0.364
St. Deviation	99.666	2.713	8.038	2.292	0.256	0.221	0.119
Median	43.455	3.155	8.432	1.814	0.617	0.339	0.357
Max	742.198	18.580	60.319	16.894	1.000	1.000	0.874
Min	-1.135	0.000	0.000	0.000	0.000	0.066	0.000
TSE (N=181)	**loglik**	**eps_b**	**mu**	**eps_s**	**delta**	**alpha**	**PIN**
Mean	135.262	5.020	14.905	4.496	0.571	0.355	0.318
St. Deviation	249.956	6.854	16.987	6.929	0.280	0.219	0.126
Median	61.747	2.884	10.423	2.535	0.557	0.316	0.314
Max	2553.183	62.831	151.026	70.158	1.000	1.000	0.779
Min	-1.724	0.000	0.000	0.000	0.000	0.016	0.000

Note: compiled by the author.

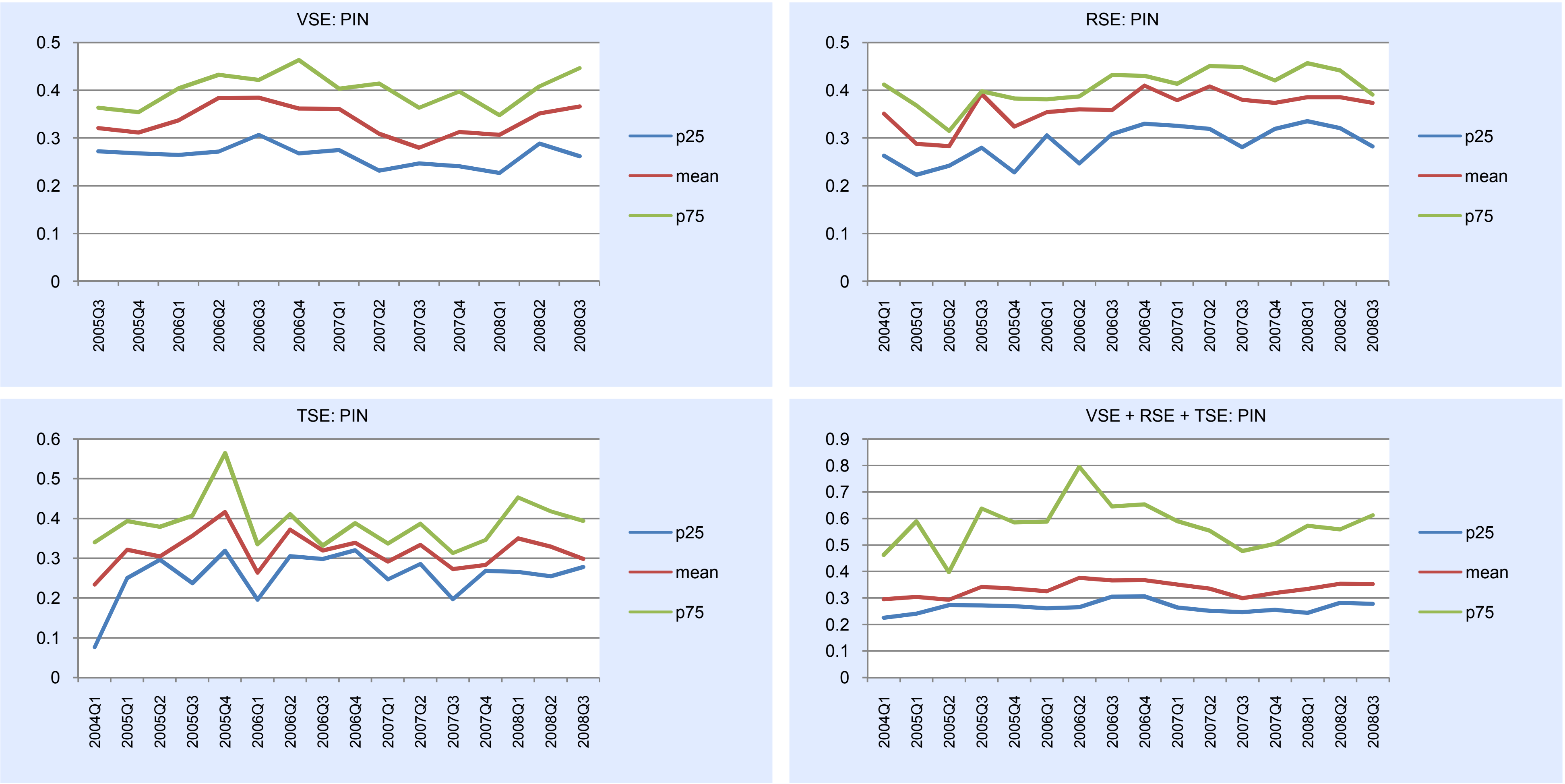
Appendix D

Figure 2. *Distribution of PIN*



Note: compiled by the author.

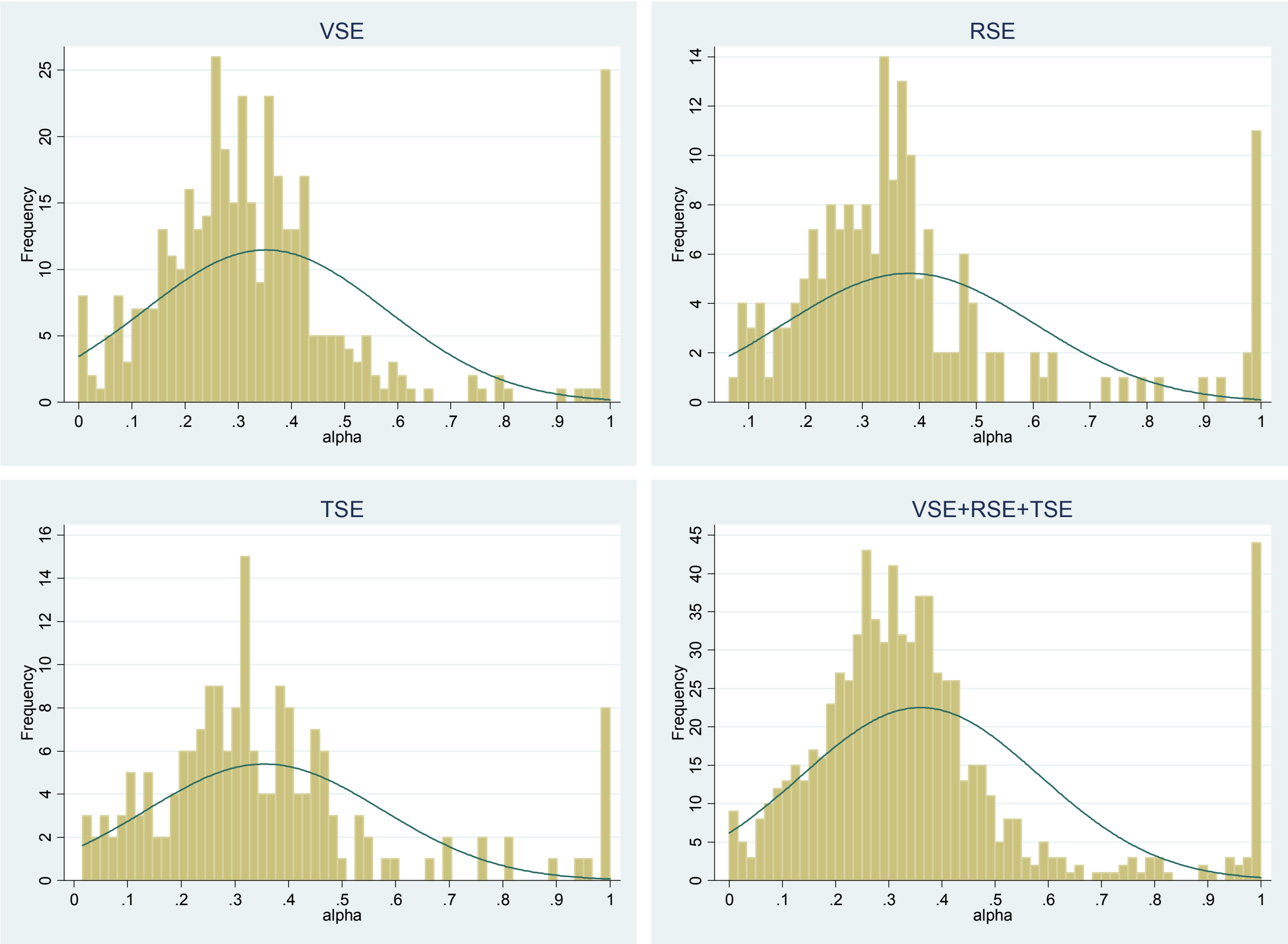
Figure 3. *25th, 75th percentiles and mean of PIN across time*



Note: compiled by the author.

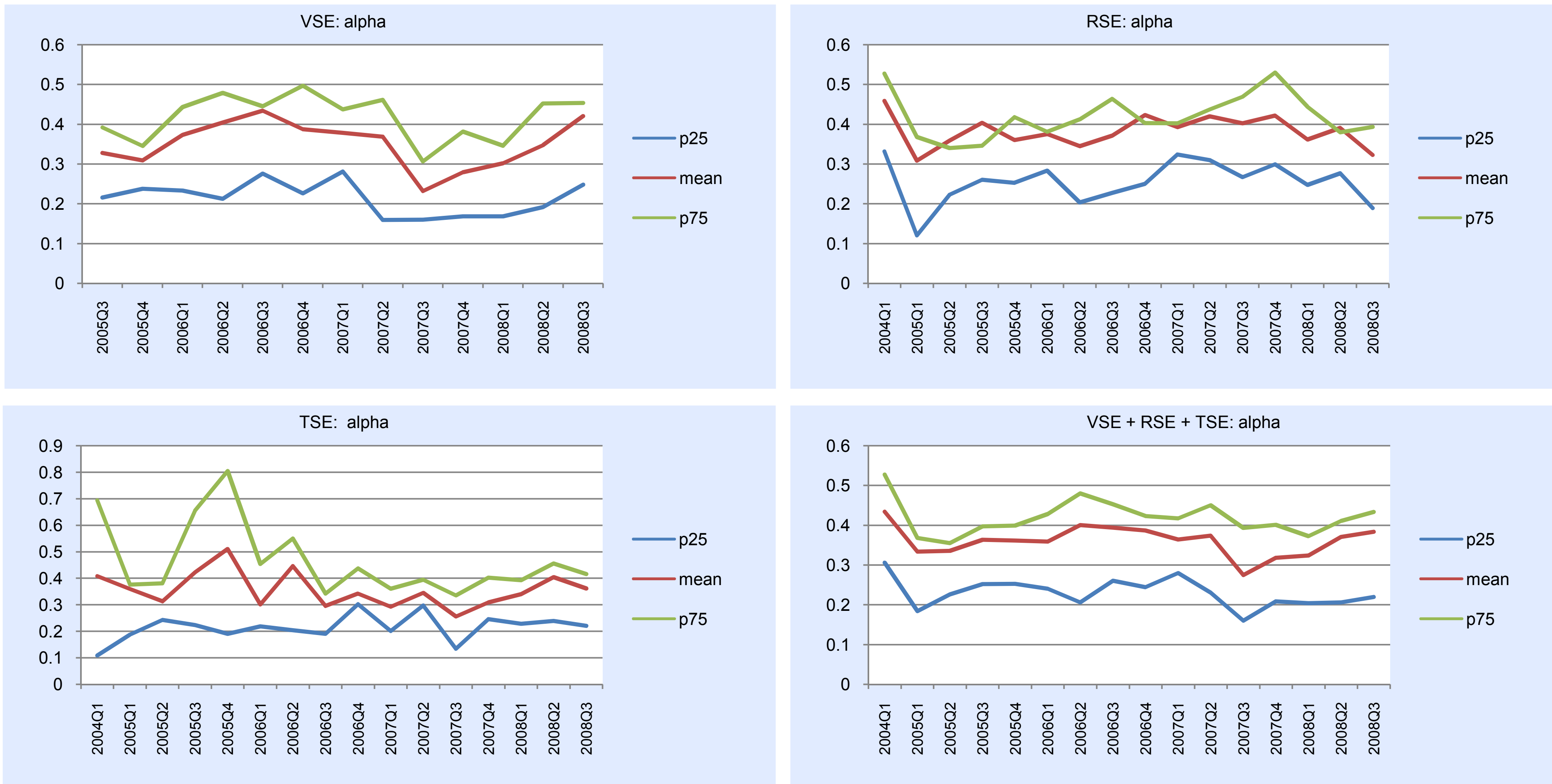
Appendix E

Figure 4. *Distribution of parameter α*



Note: compiled by the author.

Figure 5. *25th, 75th percentiles and mean of α across time*



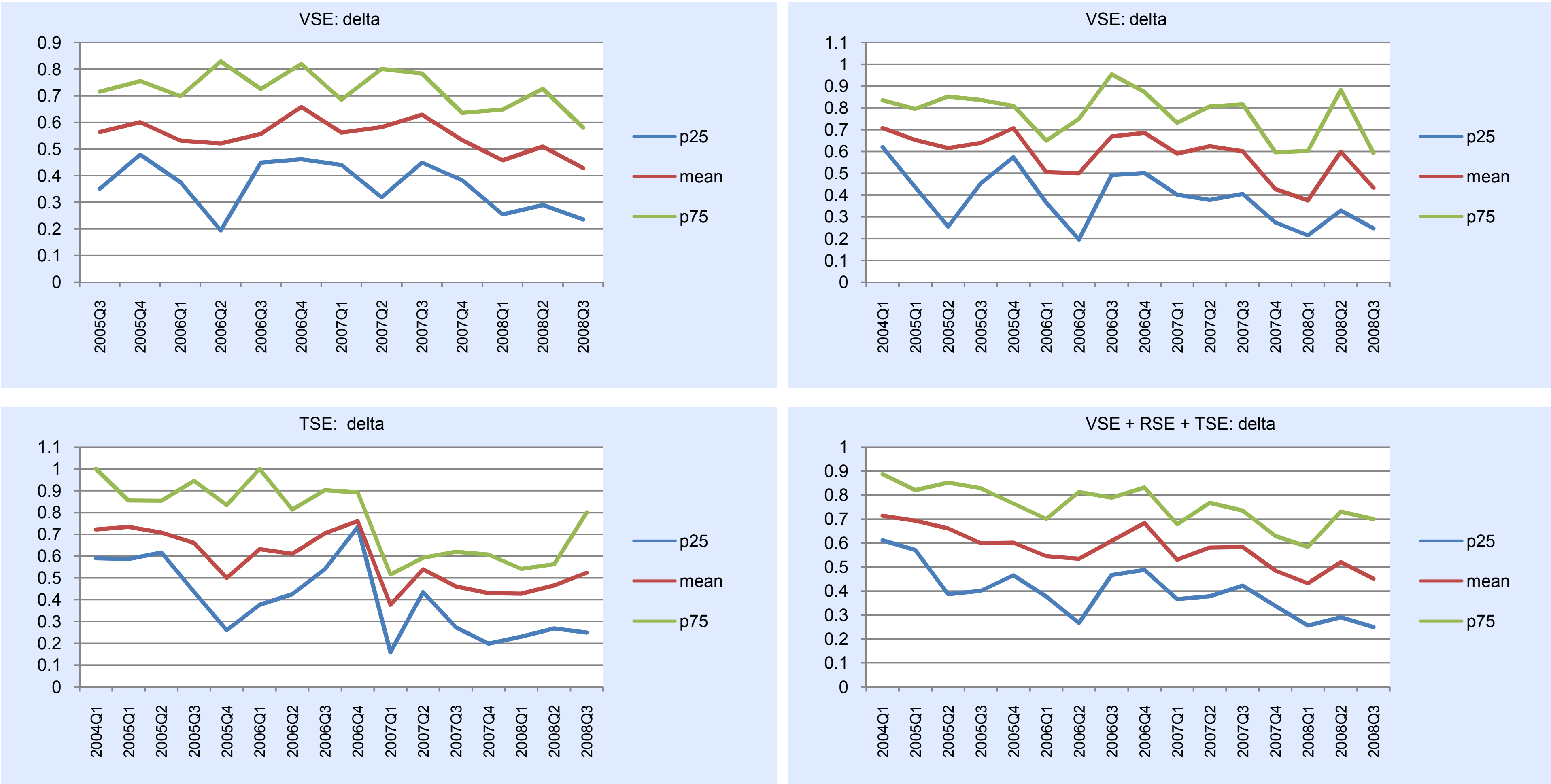
Note: compiled by the author.

Appendix F

Figure 6. *Distribution of parameter δ*

Note: compiled by the author.

Figure 7. *25th, 75th percentiles and mean of δ across time*



Note: compiled by the author.

Appendix G

Table 3

Summary statistics of asset pricing variables

Pooled data (N=759)	R	β	SIZE	BM
Mean	-0.047	0.84	191,486,690	0.95
St. Deviation	0.251	0.53	303,617,063	1.01
Median	-0.042	0.77	76,719,578	0.65
Max	0.895	2.75	2,618,530,984	9.22
Min	-0.913	-1.14	835,461	0.10
VSE (N=392)	**R**	**β**	**SIZE**	**BM**
Mean	-0.066	0.71	212,923,266	0.97
St. Deviation	0.238	0.49	348,593,750	1.06
Median	-0.058	0.68	85,797,034	0.65
Max	0.873	2.07	2,618,530,984	9.22
Min	-0.913	-1.14	835,461	0.11
RSE (N=186)	**R**	**β**	**SIZE**	**BM**
Mean	-0.043	1.09	125,723,796	1.34
St. Deviation	0.230	0.52	156,249,071	1.16
Median	-0.029	1.06	54,582,155	1.03
Max	0.703	2.75	596,112,145	7.77
Min	-0.727	0.16	1,895,265	0.25
TSE (N=181)	**R**	**β**	**SIZE**	**BM**
Mean	-0.012	0.89	212,640,064	0.51
St. Deviation	0.292	0.54	305,318,558	0.41
Median	-0.033	0.78	84,399,998	0.45
Max	0.895	1.97	1,158,818,009	3.83
Min	-0.800	-0.78	3,309,669	0.10

Note: compiled by the author.

Table 4

Summary statistics of portfolios PIN Low and PIN High

Pooled data (N=759)	Number of observations	Average	Std. Err.	[95% Conf. Interval]	
R					
PIN low	377	-0.052	0.012	-0.076	-0.028
PIN High	382	-0.025	0.023	-0.069	0.020
β					
PIN low	377	0.885	0.027	0.832	0.939
PIN High	382	0.804	0.027	0.750	0.858
SIZE					
PIN low	377	18.425	0.073	18.282	18.568
PIN High	382	17.876	0.075	17.730	18.023
BM					
PIN low	377	-0.495	0.038	-0.570	-0.421
PIN High	382	-0.280	0.042	-0.364	-0.197

Appendix H

Table 5

Quarterly correlations of PIN and turnover

			2004Q1
			-0.11
2005Q1	**2005Q2**	**2005Q3**	**2005Q4**
0.04	-0.17	-0.10	-0.17
2006Q1	**2006Q2**	**2006Q3**	**2006Q4**
-0.13	-0.26	-0.28	-0.11
2007Q1	**2007Q2**	**2007Q3**	**2007Q4**
-0.25	-0.08	-0.16	-0.28
2008Q1	**2008Q2**	**2008Q3**	
-0.23	-0.13	-0.14	

Table 6

Correlations of arrival rates of traders and turnover

	eps_b	**eps_s**	**mu**
Turnover	0.52	0.50	0.51

Table 7

Correlation matrix

Pooled data (N=759)	**R**	**β**	**PIN**	**SIZE**	**BM**
R	1				
Sig. level					
β	-0.187	1			
Sig. level	0.000				
PIN	-0.006	-0.060	1		
Sig. level	0.871	0.098			
SIZE	0.001	0.161	-0.160	1	
Sig. level	0.974	0.000	0.000		
BM	-0.071	0.154	0.153	-0.397	1
Sig. level	0.051	0.000	0.000	0.000	

Note: compiled by the author.

Appendix J

Table 8

Regression analysis

The table contains regressions results based on time-series averages of cross-sectional asset pricing tests using Fama-MacBeth two step procedure (1973) under Litzenberger and Ramaswamy (1979) precision weighted adjustment (WLS). The dependent variable is an excess return of stock *i* on quarter *t*. Coefficient β is an equity beta estimate, PIN is the log of lagged probability of information-based trading, *SIZE* is a log of firm's lagged capitalization in EUR, and *BM* is a log of company's quarterly lagged book equity value divided by lagged quarterly market equity value.

Panel A: VSE

R	coef.	t	P>t
beta	-0.0931	-3.36	0.003
PIN	0.0038	1.43	0.089
SIZE	0.0233	3.50	0.002
BM	0.0102	2.91	0.006
_cons	-0.4250	-3.55	0.002

Panel B: RSE

R	coef.	t	P>t
beta	-0.0420	-3.83	0.001
PIN	0.0020	0.66	0.259
SIZE	0.0147	3.75	0.001
BM	0.0457	3.75	0.001
_cons	-0.3602	-3.66	0.001

Panel C: TSE

R	coef.	t	P>t
beta	-0.1082	-3.40	0.002
PIN	0.0001	0.34	0.370
SIZE	0.0005	0.44	0.332
BM	0.0146	3.23	0.003
_cons	0.0614	1.74	0.052

Panel D: VSE + RSE + TSE

R	coef.	t	P>t
beta	-0.0630	-3.32	0.002
PIN	0.0001	0.39	0.351
SIZE	0.0150	3.71	0.001
BM	0.0032	1.72	0.053
_cons	-0.2577	-3.68	0.001

Note: compiled by the author.