Iranian Journal of Economic Studies, 12(1) 2023, 31-49

Iranian Journal of Economic Studies

Journal homepage: ijes.shirazu.ac.ir

How the asymmetric information affects the Stock Returns: Approach of VPIN on the Tehran Stock Exchange

Neda Esmaeeli[a]*, Maryam Ihami[a]

a .Department of Applied Mathematics and Computer Science,, University of Isfahan, Isfahan, Iran.

Article History

Received date: 06 June 2023
Revised date: 04August 2024
Accepted date: 13 September 2024
Available online: 19 October 2024

JEL Classification

G12
G14
D4
D82

Keyword
Asymmetric Information
Stock Returns
Information Risk
Multi-Factor Asset Pricing Models
Volume-Synchronized Probability of Informed Trading (VPIN)

Abstract

This paper aims to assess the existence of information asymmetry on the Iranian stock market and its impact on expected portfolio returns by applying Volume-Synchronized Probability of Informed Trading (VPIN) as a measurement tool. To this end, we used the actual data of 40 companies on the Tehran Stock Exchange (TSE) within the period from March 22, 2018 to March 19, 2020. The outcomes highlight the presence of moderate toxicity levels in the orders of these stocks. Since asymmetric information leads to a risk to investors, they may ask for a premium to trade riskier assets based on the information level, which means that market makers may incorporate the information risk into the pricing of assets. To check this, we investigated the effect of asymmetric information risk on the stock returns on the TSE by adding a factor about the level of order toxicity to the 3, 4, and 5-factor asset pricing models. According to our findings, we affirm that the Iranian stock market priced the asymmetric information risk during the time interval from March 22, 2018, to March 19, 2020. Therefore, it is essential to take into account the information risk factor besides a combination of factors such as market, size, profitability, and investment to obtain the most efficient explanation for the returns of portfolios.

Highlights

- Asymmetric information has crucial and non-negligible impacts on financial markets, and we can help increase market transparency by quantifying and measuring it.
- By applying VPIN as a measurement tool of information asymmetry, the findings suggest the presence of moderate levels of stock order toxicity on the Tehran Stock Exchange (TSE) within the time interval of the study.
- The findings also affirm that the Iranian stock market priced the asymmetric information risk within the time interval of the study.

* n.esmaeeli@sci.ui.ac.ir
DOI: 10.22099/ijes.2024.47570.1909

1. Introduction

The financial markets play an important role as one of the main mainstays of the economy. One of the most important concerns of policymakers in these markets is to realize necessary conditions for the investors who intend to obtain the maximum return on their investments with the least risk, and it is obvious that information plays an essential role in the meantime. Asymmetric information in financial markets refers to a situation in which one side involved in a financial deal has more information than the other hand and can make informed decisions. Information asymmetry significantly affects price levels, market formation, and investment risk interactions. In other words, asymmetric information can lead to market inefficiencies, such as mispriced securities or inefficient allocation of resources. Measuring and understanding the extent of asymmetric information can help investors and market participants identify and take advantage of these inefficiencies. Furthermore, measuring asymmetric information can help increase market transparency. This can lead to the implementation of regulations or disclosure requirements aimed at leveling the playing field for all market participants. Moreover, understanding the presence and extent of asymmetric information can help financial institutions and investors better assess and manage their risks. This can lead to the development of better risk management strategies and a more resilient financial system. On the other hand, asymmetric information can cause moral hazard and adverse selection issues, where one side takes advantage of their superior information to the detriment of others. Measuring and addressing these issues can help diminish the negative impact of asymmetric information on the market stability and efficiency.

Overall, since many investment decisions in financial markets are driven by the level of asymmetric information, measuring asymmetric information in a financial market can contribute to a more efficient, transparent, and stable market environment, benefiting investors, market participants, and the economy as a whole. Hence, the issue of asymmetric information has gained significant attention recently, leading to the development of various mathematical models to address financial markets with different levels of information. It is worth noting that in recent years and based on experimental observations, there have been claims in the real world regarding the trade of stocks based on insider information, including some cases in international stock exchanges, which have highlighted suspicious transactions by some companies on the verge of financial crisis. On the Tehran Stock Exchange, stock symbols such as "Rkish", "Akntor", "Saipa", "Khodro", and "Zob" can also be pointed out as examples of suspicious transactions (Taleblo, R., Rahmaniani, M. 2017). Since the asymmetric information in the market is not directly observable, researchers have used some criteria to measure it. One of the newest of these metrics is Volume-Synchronized Probability of Informed Trading (VPIN) which attempts to directly quantify the toxicity level present within the order flow of a particular stock. We calculated VPIN for the selected companies in TSE over the period spanning from March 21, 2018, to March 19, 2020. The outcomes highlight the presence of moderate toxicity levels in the orders of these stocks. We further analyzed our list of companies by dividing them into two distinct segments based on their size. Subsequently, we calculated VPIN independently for each segment. In line with market expectations, the shares of large companies are more widely attended by analysts and investors. Hence, experiencing a privileged transaction is theoretically less feasible for these shares, leading to a lower VPIN compared to shares of smaller firms. Our segment analysis results have indeed validated this expectation. However, the results do not indicate that the toxicity level of the stocks from the companies that belong to the theoretically more overt segments of the market is significantly lower. As mentioned earlier, an asymmetry of information regarding assets being exchanged within a financial market can introduce a potential risk for investors. Consequently, investors might claim a higher premium to engage in the trading of assets that are deemed to carry a higher level of risk. Therefore, one of the factors that market makers take into account when determining the price of an asset is the level of information risk it carries. This paper sought to assess whether stock order flow's toxicity, measured by VPIN, is a systematic risk factor that is taken into account by investors when pricing shares on the Tehran Stock Exchange (TSE) within the period from March 21, 2018, to March 19, 2020. This paper also investigates the explanatory power of information risk besides the pre-known factors such as market, size, profitability, and investment, concerning the cross-sectional study of expected stock returns on TSE. The results show that the asymmetric information risk is priced in TSE and adding this additional factor related to the information risk into the 3, 4, and 5-factor asset pricing models results in a stronger explanatory power for these models.

2. A Review of the Related Literature

Over the last few years, several studies on stock returns and the effect of asymmetric information on the Iranian economy, especially in financial markets, have been conducted. See, for instance: Mahinizadeh et al. (2018), Cheshomi & Osmani (2021), Miri & Kiani (2015) where the effect of the stock market on Iran's economic growth has been investigated under the condition of information asymmetry, Molaei et al. (2016) where an experimental analysis was used to study the potential for jumps in the index price of large companies on the Iranian stock market as a sign that new information exists in the market. Furthermore, measuring the level of information asymmetry by the probability of informed trading (PIN) and its effects on the daily returns of selected companies in the TSE was investigated in Shamsoddini et al. (2016) and Taleblo & Rahmaniani (2017). In Shamsoddini et al. (2016), the authors analyzed information asymmetry in larger firms in the TSE using the PIN and FE indices. Due to incompatible findings, they proposed a composed index to better capture this issue in developing markets. Their study indicate that this new combinatory index yields superior outcomes, combining aspects of PIN and FE. In Taleblo and Rahmaniani (2017), PIN was used as an indicator of market information asymmetry for 12 selected companies on TSE. Parameter estimation was conducted using maximum likelihood with the R package. Finally, the symbols of the largest firms indexed on the TSE are classified in Mirbagherijam (2020) based on PIN. In Mirbagherijam (2020), the trading symbols of the 30 largest TSE companies were ranked using the PIN index from March 20, 2015, to March 19, 2017, along with a modified clustering algorithm (EA). In addition, he utilized the analysis of variance (ANOVA) technique to identify the factors contributing to variations in the estimated PIN. Moreover, in Mirbagherijam (2020), it was affirmed that weekdays did not significantly impact the PIN index.

Among the well-known studies abroad, in Abad et al. (2012), the authors employed PIN and VPIN to evaluate the risk of information asymmetry on the Spanish stock market. Furthermore, the problems of insider trading, order flow toxicity, and its effects on daily transactions of the Australian stock market were considered in Wei et al. (2013). The role of high-frequency trading in order flow toxicity and stock price variance on the U.S. markets was investigated in Van Ness et al. (2017). Using VPIN as a metric to measure the level of toxicity, they demonstrate a negative correlation between the order flow toxicity level and high-frequency trading. The last, among others, the impact of asymmetric information risk on the stock market in Brazil was studied in Siqueira, et al. (2017). In their study, various statistical models have been considered to investigate this effect on stock returns, and their results show the high level of order toxicity on the Brazilian stock market.

3. The Study Model

3.1 How to measure asymmetric information

Today, the level of asymmetric information is measured by sequential trading models based on the order flow toxicity in the microstructure of financial markets. PIN was initially presented in Easley et al. (1996). As a development of PIN, a new metric called VPIN was extended in Easley et al. (2012).

3.1.1 PIN

In 1996, a microstructure measure of information asymmetry known as PIN was presented in Easley et al. (1996). The authors suggested a market microstructure theory that explains the domain at which price makers are ready to supply liquidity, based on a sequent trading model with Bayesian updates. This theory has been widely accepted among academics and practitioners and has been appended in most market microstructure course books. In their model, the number of buyers and sellers who started trading with and without extra information within any trading day is considered to be realizations of independent Poisson distributions. Furthermore, the model assumes that the mean of this distribution depends on the fact that no news, good news, or bad news arrived during that day. Let denote the probability of news arrival by α. Considering that there is news, δ is the probability that a security price will be negatively affected by the news event, and accordingly $(1-\delta)$ is the probability of a positive effect of the news event. Another consideration is that the liquidity traders are available on the market during a given trading period to either purchase or sell the asset. The traders with additional information await a noxious impact on the price of the security as soon as bad news comes up. Therefore, they likely intend to sell the security. It is assumed that order arrivals during the bad news period have an independent Poisson distribution with means mu for the informed dealers, and for the liquidity buy and sell dealers respectively. Then, the probability of negotiation with private information is $PIN = \frac{\alpha\,\mu}{\alpha\,\mu+\epsilon_b+\epsilon_s}$. The PIN model requires an estimate of five non-observable parameters α, $\mu, \epsilon_b, \epsilon_s$, δ. These parameters can be estimated by maximizing a likelihood function.

3.1.2 VPIN

In 2012, VPIN was introduced in Easley et al (2012). A volume clock is used in this method to keep the data sampling in sync with the market activity that is recorded by regular volume buckets. This method begins by splitting the sample of time bars into volume buckets, where each group consists of a set of trades that is identical in volume. Let S represent the price process of security. In a volume bucket, V_τ^B represents the volume amount categorized as buy and is calculated as follows:

$$V_\tau^B = \sum_{i=t(\tau-1)+1}^{t(\tau)} V_i N\left(\frac{S_i - S_{i-1}}{\sigma_{\Delta S}}\right) \tag{1}$$

where *t(τ)* is the index of the final bar which is involved in bucket τ, V_τ^B is the buy volume, V_i is the complete volume for every bucket, N (.) is the standard normal distribution function, and $\sigma_{\Delta S}$ is the standard deviation of price variations among the bars. Moreover, we have

$$V_\tau^S = \sum_{i=t(\tau-1)+1}^{t(\tau)} V_i \left(1 - N\left(\frac{S_i - S_{i-1}}{\sigma_{\Delta S}}\right)\right) \tag{2}$$

Since all buckets have the same amount of volume V, we get

$$\frac{1}{n}\sum_{\tau=1}^{n}(V_\tau^B + V_\tau^S) = V = \alpha\mu + 2\epsilon \tag{3}$$

where *n* is the number of volume buckets employed to compute VPIN. It can be proved that VPIN is an acceptable approximation of PIN with,

$$VPIN = \frac{\sum_{\tau=1}^{n}|V_\tau^S - V_\tau^B|}{nV} \tag{4}$$

because $E\,[|V^S - V^B|] \approx \alpha\mu$, and $PIN = \frac{\alpha\mu}{\alpha\mu+2\epsilon}$, where $\epsilon_b = \epsilon_s = \epsilon$.

3.2 Asset pricing models

One of the earliest and most remarkable asset pricing models is the Capital Asset Pricing Model (CAPM), which was extended in Sharpe (1964). The CAPM attractiveness stems from the hypothesis that the expected return of assets and its beta are linearly related. However, the model has been subject to significant criticism over recent years, and several asset pricing models are proposed in the literature to improve it. One of these important models is the well-known 3-factor model introduced in Fama & French (1993) which adds, besides beta, two more explanatory variables to the CAPM. These factors are relevant to the book-to-market ratio and firm size. In Fama & French (2015), the authors introduced their 5-factor model by considering two factors of investment and profitability to the 3-factor model. The profitability factor displays the difference in return between the portfolio of the most profitable stocks and the portfolio of the least profitable stocks on the market, and the investment factor describes the return difference between the portfolios of the most conservative stock and the most aggressive stocks respectively. In Carhart (1997), a 4-factor model was introduced by adding a cross-sectional momentum factor for asset pricing of stocks, which explains the persistence of high returns for mutual funds. The following table represents the 3, 4, and 5-factor models to examine their characteristics:

Table 1. Multifactor models for stock returns

Model	Model specifications
3-Factor	$R_{it} - R_{ft} = \alpha_i + \beta_i MKT_t + s_i SMB_t + h_i HML_t + \varepsilon_{it}$
4-Factor	$R_{it} - R_{ft} = \alpha_i + \beta_i MKT_t + s_i SMB_t + h_i HML_t + u_i UMD_t + \varepsilon_{it}$
5-Factor	$R_{it} - R_{ft} = \alpha_i + \beta_i MKT_t + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + \varepsilon_{it}$

Source*: Fama and French (1993, 2015) and Carhart (1997).*

In the above table , R_{it} denotes the return of portfolio i $(i = 1,\dots,n)$ at time t $(t = 1,\dots, T)$ and R_{ft} is the risk-free rate of return at time t Furthermore, $R_{it} - R_{ft}$ is a dependent variable, which is called the excess return of portfolio i at time t. MKT_t[1], SMB_t[2], HML_t[3], UMD_t[4], RMW_t[5], and CMA_t[6] are independent variables stand respectively for the return of the market portfolio, the size factor, the book-to-market factor, the momentum factor, the profitability factor, and the investment factor at time t, and they are defined as follows:

- MKT is the return spread between the capitalization-weighted stock market and cash.
- SMB is the return spread between small and large stocks.
- HML is the return spread between inexpensive and expensive stocks.
- UMD is the return spread between the lowest-performing and highest-performing stocks.
- RMW is the return spread between the most and least profitable firms.
- CMA is the return spread between firms that invest conservatively and those that invest aggressively.

[1] Market (MKT is an abbreviation for the market risk premium factor)
[2] Small Minus Big
[3] High Minus Low
[4] Up Minus Down
[5] Robust Minus Weak
[6] Conservative Minus Aggressive

4. Methodology

4.1 Hypothesis

Two major propositions are investigated as follows:

First: The toxicity level of order flows increases as the size of the company decreases.

Second: Portfolio returns are better explained by a factor linked to VPIN.

4.2 Data collection

The data used in this study include the stock prices of Tehran Stock Exchange (TSE) companies along with their financial statements, short-term deposit interest rate calculated by the mean of the monthly deposit interest rate from March 22, 2018, to March 19, 2020[1], as the risk-free rate of return, and finally Tehran Exchange Dividend and Price Index (TEDPIX) as a criterion for calculating the market return. The statistical population of this study is all companies listed on TSE, which we have selected from March 22, 2018, to March 19, 2020, with the following restrictions:

1. Companies whose fiscal year ends on March 26 and have no change in their fiscal year during the review period.
2. Companies should not be involved in financial intermediation.
3. Companies should not have a trading break of more than three months.
4. The book value of companies should be positive during the period under review.

4.3 Multifactor models and VPIN

In Siqueira et al. (2017), generalized versions of the 3, 4, and 5-factor models are tested by considering an additional factor called IMU[2]. This factor, which measures the order flow toxicity, has been subjoined to the previous multifactorial models as an independent variable to study the impact of asymmetric information on the return deviations. In other words, IMU is the return spread of firms with a high level of VPIN minus a low level of VPIN. The following table presents these models:

Table 2: Multifactor models for the stock returns plus the factor IMU

Model	Model specifications
3-factor + IMU	$R_{it} - R_{ft} = \alpha_i + \beta_i MKT_t + s_i SMB_t + h_i HML_t + \rho_i IMU_t + \varepsilon_{it}$
4 -factor + IMU	$R_{it} - R_{ft} = \alpha_i + \beta_i MKT_t + s_i SMB_t + h_i HML_t + u_i UMD_t + \rho_i IMU_t + \varepsilon_{it}$
5-factor + IMU	$R_{it} - R_{ft} = \alpha_i + \beta_i MKT_t + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + \rho_i IMU_t + \varepsilon_{it}$

Source: *Siqueira et al. (2017).*

This study aims to investigate these generalized models for the stock returns of a group of selected companies in the Iranian stock market.

---

[1] www.cbi.ir

[2] Informed Minus Uninformed

4.4 Dependent and independent variables

Dependent variables comprise the average of monthly excess return on the stock portfolios formed based on the method pursued in Fama & French (2015), and independent variables include MKT_t, SMB_t, HML_t, UMD_t, RMW_t, CMA_t, and *IMUt*.

4. 5 Portfolio Formation

To calculate the independent variables, we first need to create our portfolios. For this purpose, we consider 2×2 portfolios and sort our companies based on their size variable into small (S) and big (B) groups, and then in each group, we have two classes based on the value factor which are high (H) and low (L). The medians of the size and the value factors are the middle values in the sorted lists, respectively. Let R stand for the average monthly return on an investment portfolio. Then after our classification, we have

$$SMB = \frac{R_{SH} + R_{SL}}{2} - \frac{R_{BH} + R_{BL}}{2}. \tag{5}$$

$$HML = \frac{R_{SH} + R_{BH}}{2} - \frac{R_{SL} + R_{BL}}{2}. \tag{6}$$

Furthermore, RMW and CMA in the 5-factor model are given by:

$$RMW = \frac{R_{SR} + R_{BR}}{2} - \frac{R_{SW} + R_{BW}}{2}. \tag{7}$$

$$CMA = \frac{R_{SC} + R_{BC}}{2} - \frac{R_{SA} + R_{BA}}{2}. \tag{8}$$

and UMD in the 5-factor model is calculated as follows:

$$UMD = \frac{R_{SU} + R_{BU}}{2} - \frac{R_{SD} + R_{BD}}{2}. \tag{9}$$

For the additional factor IMU, we first calculated VPIN for each company every month, then we divided the companies into two groups hinge on the median of this factor. The classification based on the factors of size and VPIN is as follows:

Table 3. Portfolios created based on the variable's size and VPIN

Portfolio	Initials	Description
Small and High	SH	Companies with small-size and high VPIN variables
Big and High	BH	Companies with big-size and high VPIN variables
Small and Low	SL	Companies with small-size and low VPIN variables
Big and Low	BL	Companies with big-size and low VPIN variables

Source: *Research finding*

Moreover, we have

$$IMU = \frac{R_{SH} + R_{BH}}{2} - \frac{R_{SL} + R_{BL}}{2}. \tag{10}$$

5. Results

5.1 Results for VPIN

Table 4 shows the results of the calculation for VPIN.

Table 4. Descriptive statistics of the VPIN average for the selected companies from April 2018 to March 2020

Number of companies	Min	Max	Average	Std. Dev.
40	0.1102	0.4238	0.2517	0.0707

Source: *Research finding*

5.2 VPIN analysis based on the market value of companies

The fact that VPIN is lower for the stocks of large firms is one of the most significant findings in studies that have looked into this issue. For instance, the investigations obtained in the Spanish, Australian, and U.S. stock markets have verified this relation (see Abad & Yagüe (2012), Wei et al. (2013), and Van Ness et al. (2017) respectively). To investigate the relevance between VPIN and the Stock returns of some companies in the Iranian stock market, we classified the stocks as small and large, based on the average of their monthly market values, and to gain a profound insight into this relevance. Table 5 shows the descriptive statistics for these two groups.

Table 5. Descriptive statistics of the VPIN average contained in each group

Descriptive statistics	Groups	
	Small	Large
Min	0.1369	0.1102
Max	0.4238	0.4234
Mean	0.2645	0.2390
Std. Dev.	0.0713	0.0696

Source: *Research finding*

The monthly VPIN behavior of small and large groups of selected companies on the Iranian stock market from April 2018 to March 2020 is depicted in Figure 1. It is observed that the calculated VPIN for stocks of the selected companies differs significantly concerning the variable size of those companies, and the small group of companies shows a high level of VPIN. Therefore, the assumption that VPIN and its size have a negative correlation was confirmed by our analyzed sample on the Iranian stock market.

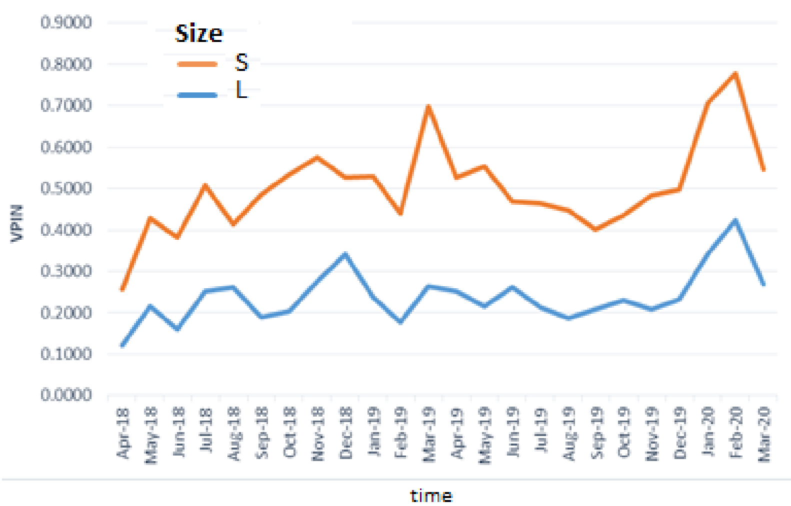

Figure 1. Monthly VPIN for the stock size groups.
S: small stock groups, L: large stock groups.
Source: *Research finding.*

5.3 Factor analysis in 3, 4, and 5-factor models reinforced with IMU

Our first step in investigating the implications of the models (presented in Table 2) is to analyze the correlation between their factors. According to the results obtained in Tables 6 and 7, we find that IMU has a weak positive correlation with the MKT factor. There is also a weak and moderate positive correlation, respectively, between IMU and HML and IMU and SMB. A negative correlation between IMU and UMD is reported in Table 6. The two factors of RMW and CMA are negatively related to IMU.

Table 6. Correlation between the 3- and 4-factor models and IMU

	MKT	SMB	HML	UMD	IMU
MKT	1				
SMB	-0.0288	1			
HML	0.1092	0.3670	1		
UMD	0.1692	-0.1700	-0.4239	1	
IMU	0.0567	0.1820	0.0877	-0.4426	1

Source: *Research finding.*

Table 7. Correlation between the 5-factor model and IMU

	MKT	SMB	HML	RMW	CMA	IMU
MKT	1					
SMB	-0.0288	1				
HML	0.1092	0.3670	1			
RMW	-0.0232	-0.4693	-0.6634	1		
CMA	-0.0937	-0.0837	-0.1018	0.4405	1	
IMU	0.0567	0.1820	0.0877	-0.3847	-0.2250	1

Source: *Research finding.*

5. 4. Factor regression analysis

To check whether our factors are not priced, we have carried out the Factor regression analysis proposed in Mohanram and Rajgopal (2009) and Fama & French (2015). In other words, when a regression intercept is equal to 0, it means that the return is accounted for by the other factors. Appendix A contains the factor model regressions results. Table 8 presents the 3-factor regression estimates and the IMU factor. The fact that the intercept coefficient of the SMB factor is statistically different from 0 indicates that other factors do not already incorporate its predictive power. Furthermore, the regression R2 has been substantially increased by adding IMU and so the independent variables have more power to explain the dependent variable. Therefore, the IMU factor is effective in the Iranian stock market.

The results of the factor regression analysis for the 4-factor model are shown in Table 9. According to Table 9, none of the regression constants is equal to zero. Furthermore, adding the IMU factor does not have much effect on increasing the amount of R2. According to the calculated P-Value, HML and IMU factors affect the UMD factor since their P-Value is less than 0.05. Moreover, the only factor that changes IMU is UMD, and the other factors do not touch it. Therefore, Table 9 shows that on the Iranian stock market, the largest P-value belongs to SMB. Specifically, its intercepts are equal to 0.65 and 0.82.

At last, in Table 10, we analyze the five factors regression estimates. Based on the obtained results in Table 10, the regression constant is not zero for any of the factors, so all the factors are relevant and are not captured by the others. Furthermore, the IMU and SMB factors have the highest probability of being fixed, respectively. According to the calculated probability, the only factor affecting CMA is RMW, and the factors HML and CMA affect RMW. The other factors do not have any significant effect on each other. Eventually, on the Iranian stock market, two factors RMW and SMB have shown a regression constant close to zero and have a high P-value.

5. 5. Statistical Significance of factors in the multifactorial models

The results of the t-test statistics for the multifactorial models are submitted in Appendix B. Table 11 contains the results for the 3-factor model. Based on the results, the significance level of the F-statistic is zero. Thus, the hypothesis that all coefficients are zero is rejected at a 95% confidence level. Moreover, based on the obtained P-Values, we can deduce that all factors affect stock returns. However, the value of the adjusted coefficient of determination (R2) equals 0.15 which implies that only 15% of the stock return changes are elucidated by the variables of the 3 and 5-factor models. Now, to verify the effect of the IMU factor, we have added this factor to the 3-factor model. The results are presented in Table 12. According to the observed P-Values, all factors, including the IMU factor, have influences on stock returns, and adding this factor has not destroyed the significance of the other factors.

Similarly, we check the effect of factors on the stock returns in the 4-factor model. The results are summarized in Table 13. According to the P-Values, it is observed that among the variables, only the HML factor does not affect the stock return. We have added the IMU factor to the 4-factor model and the results of the regression analysis are displayed in Table 14. The obtained results represent that the null hypothesis has been rejected at a 95% confidence level. Furthermore, the P-Values show that SMB, MKT, and IMU affect the stock return while HML and UMD are non-effective. Moreover, adding the IMU factor to the 4-factor model made the UMD factor ineffective. Finally, we examine how the factors affect the stock returns under the 5-factor model. Table 15 provides a short of the calculated results. According to the P-Values in Table 15, HML and CMA do not affect the stock returns. The IMU factor is also added to this model, and the regression results after adding the IMU factor are presented in Table 16. Based on the P-Values, at a 95% confidence level, HML and CMA factors do not change the stock returns while the others are effective.

6. Concluding Remarks

This paper inquiries about the problem of asymmetric information on the Iranian stock market. More specifically, our goal is to measure the toxicity level of stocks on the Tehran Stock Exchange (TSE) and to explore if TSE investors are pricing the information risk. For this purpose, VPIN has been used to measure the toxicity of the stock orders for 40 selected active companies in the Iranian stock market from 2018 to 2020. The average of VPIN for these companies is equal to 0.2517 and the standard deviation is 0.0707.

The selected companies were divided into small and large groups in terms of their market value. The results indicate a slight decline in the VPIN value of companies based on an increase in their size. For the second objective, the impact of asymmetric information on the stock return deviations of these 40 chosen companies was also examined. To do this, we employed VPIN as a well-known criterion to calculate the risk of asymmetric information and added it through a new factor called IMU (presented in Siqueira, et al. (2017)) to the multifactor models introduced in Table 1. We evaluated the effectiveness of the IMU factor through these multifactor models and the obtained results assert that the information risk has a direct impact on the stock returns of the selected companies.

Furthermore, we studied the effect of the firm size, the company value (ratio of the book value to the market value), the market risk expenditure, the profitability, the investment, the tendency to past stock performance, and the information risk factor on the returns of the selected stocks. Overall, the factors of size, market risk, and stock returns in all the models are affected by asymmetric information risk. In the 3-factor model, the value factor was also significant. However, the value factor was unneeded in the 4-factor model. Moreover, adding the IMU factor made the factor of tendency to past performance irrelevant. In the 5-factor model, the value and investment factors were not significant.

This study is the first to apply VPIN as a metric to assess the toxicity level of order flows on the Tehran Stock Exchange and as a factor in analyzing the impact of asymmetric information on stock returns in the market. Previous studies in Shamsoddini et al. (2016) and Taleblo & Rahmaniani (2017) used PIN to measure the toxicity level of order flows on TSE. In Shamsoddini et al. (2016), information asymmetry in large firms on the TSE is measured by PIN, FE, and C-PIN-FE metrics. The PIN index showed an asymmetry of over 20%. The FE values showed significant discrepancies between earnings forecasts and actual earnings, sometimes exceeding 1000%. To address this, they introduced C-PIN-FE, a composed index integrating PIN and FE. They asserted that the C-PIN-FE index better captures information asymmetry in dynamic stock markets. In Taleblo & Rahmaniani (2017), PIN was estimated for 12 companies in TSE using maximum likelihood. Their findings indicate that the average PIN ranges between 0.35 and 0.4 across the various companies. Mirbagherijam (2020) ranked the 30 largest companies in the TSE based on their asymmetric information risk over two years from March 20, 2015, to March 19, 2017. Using a modified clustering algorithm (EA), he estimated that 88.2% of trading days had a PIN below 0.1, with 60% of those days registering a PIN of zero.

Among the studies conducted abroad, in Abad et al. (2012), PIN and VPIN were utilized to assess information asymmetry risk on the Spanish stock market. Also, in Wei et al. (2013), issues like insider trading, order flow toxicity, and their impact on daily transactions on the Australian stock market were explored. In Van Ness et al. (2017), the influence of high-frequency trading on order flow toxicity and stock price fluctuation in U.S. markets was investigated. By using VPIN as a measure of toxicity, all of them revealed an inverse relationship between order flow toxicity and high-frequency trading. In Siqueira et al. (2017), transaction data from 142 stocks on the Brazilian Securities, Commodities and Futures Exchange (BM&FBOVESPA) between May 2014 and May 2016 was studied. Results show high flow toxicity in the stock orders. In their study of the Brazilian market, no evidence was found that transparent segments have lower toxicity. The authors also examined the impact of asymmetric information risk on stock returns. They found that the inclusion of the information risk factor complements the size factor and enhances the performance of the models, which suggests that information risk may have explanatory power over portfolio returns.

All in all, by demonstrating the presence of information asymmetry risk on the Iranian stock market, these findings have significant implications for both investors and policymakers. Investors can use this information to make more informed investment decisions, while policymakers can implement measures to reduce information asymmetry and promote a more transparent and efficient market. For instance, policymakers may implement transactional constraints, such as some particular limits on both trading volumes and price ranges, to lessen information asymmetry and foster trust among investors. Therefore, policymakers will benefit from the findings of the current study in the effective implementation of trade restrictions.

Author Contributions:
Conceptualization, N.Esmaeeli; methodology, N.Esmaeeli; validation, N.Esmaeeli; formal analysis, all authors; resources, all authors; writing—original draft preparation, M.Ihami; writing—review and editing, N.Esmaeeli; supervision, N.Esmaeeli. All authors have read and agreed to the published version of the manuscript.

Funding
This research received no external funding.

Conflicts of Interest:
The authors declare no conflict of interest.

Data Availability Statement:
The data used in the study were taken from https://www.tsetmc.com and https://www.cbi.ir.

Acknowledgments
Not applicable

References

Abad, D., & Yagüe, J. (2012). From PIN to VPIN: An introduction to order flow toxicity. *The Spanish Review of Financial Economics (SRFE)*, 2(10), 74–83.

Carhart, M. (1997). On persistence in mutual fund performance, *The Journal of Finance*, 52(1), 57–82.

Cheshomi, A. & Osmani, F. (2021). Stock Market Returns in Iran in Three Waves of COVID-19 Pandemic: Evidence of Multiple Breaks Regression, *Iranian Journal of Economic Studies (IJES)*, 10(2), 339-364.

Easley, D., Kiefer, N., O'Hara, M., & Paperman, J. (1996). Liquidity, information, and infrequently traded stocks. *The Journal of Finance*, 51(4), 1405–1436.

Easley, D., López de Prado, M., & O'Hara, M. (2012). The exchange of flow toxicity.*The Journal of Trading*, 6(2), 8–13.

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. Econ. *The Journal of Finance*, 33, 383–417.

Fama, E., & French, K. (2015). A five-factor asset pricing model. Econ. *The Journal of Finance*, 116, 1–22.

Mahinizadeh, M., Yavarib, K., Valibeigic, H., & Shafieia, A. (2018). The Effect of Asymmetric Information in the Money Market on Investment, Employment, and Production in Iran, *Iranian Journal of Economic Studies (IJES)*, 7(2), 191-217.

Mirbagherijam, M. (2020). Ranking the Trading Symbols of the Largest Companies Listed in the Tehran Stock Exchange Based on the Probability of Informed Trade Criteria, *The Iranian Economic Review*, 24(3), 567–589.

Miri, S., & Kiani, H. (2015). The Impact of Asymmetric Information and Stock Market Development on Economic Growth in Iran. *Applied Economics*, 5, 43–59.

Mohanram, P., & Rajgopal, S. (2009). Is PIN priced risk? *Journal of Accounting and Economics*, 47, 226–243.

Molaei, S., & Samadi, S. (2016). Empirical Analysis of Price Jump and Asymmetric Information in Tehran Stock Exchange. *Journal of Financial Management Strategy (JFM)*, 4(13), 65–81.

Shamsoddini, M., Shahiki Tash, M. N., & Khodadad-Kashi, F. (2016). Measuring information asymmetry in large active firms on the Tehran stock exchange, Scientific Annals of Economics and Business, 63(3), 333–346.

Siqueira, L. S., Amaral, H. F, & Correia, L. F. (2017). The effect of asymmetric information risk on return of stocks traded on the BM&FBOVESPA, *Revista Contabilidade e Financas*, 28, 425–444.

Taleblo, R., & Rahmaniani, M. (2017). Measuring Probability of Informed Trading in Tehran Stock Exchange. *Journal of Economic Modeling Research* (JEMR), 8(29), 73–98.

Wei, W., Gerace, D., & Frino, A. (2013). Informed trading, flow toxicity, and the impact on the intraday trading factors. *Australasian Accounting Business and Finance Journal*, 7(2), 3–24.

Van Ness, B. F., Van Ness, R. A., & Yildiz, S. (2017). The role of HFTs in order flow toxicity and stock price variance, and predicting changes in HFTs liquidity provisions, *Journal of Economics and Finance*, 41(4), 739–762.

Appendices

Appendix A: Tables for factor model regressions

Table 8. 3-factor regression and the IMU

Variable	α	MKT	SMB	HML	IMU	R2
MKT C.	-0.0343		-0.0709	0.2350		0.0174
MKT PV.	(0.1160)		(0.7356)	(0.5582)		
MKT C.	-0.0349		-0.0803	0.2324	0.0758	0.0210
MKT PV.	(0.1201)		(0.7125)	(0.5716)	(0.7894)	
SMB C.	0.0113	-0.0781		0.7136		0.1395
SMB PV.	(0.6312)	(0.7356)		(0.0800)		
SMB C.	0.0087	-0.0864		0.6891	0.2152	0.1631
SMB PV.	(0.7164)	(0.7125)		(0.0951)	(0.4617)	
HML C.	0.0224	0.0705	0.1945			0.1491
HML P.	(0.0567)	(0.5582)	(0.0800)			
HML C.	0.0223	0.0700	0.1930		0.0110	0.1493
HML PV.	(0.0670)	(0.5716)	(0.0951)		(0.9436)	
IMU C.	0.0105	0.0482	0.1273	0.0233		0.0372
IMU PV.	(0.5683)	(0.7892)	(0.4617)	(0.9436)		

C.: *Coefficient.* ***PV.:*** *P-Value.* ***R2:*** *R-squared.*
Source: *Research finding.*

Table 9. 4-factor regression and the IMU

Variable	α	MKT	SMB	HML	UMD	IMU	R2
MKT C.	-0.04		-0.07	0.42	0.35		0.07
MKT PV.	(0.09)		(0.75)	(0.34)	(0.28)		
MKT C.	-0.04		-0.10	0.49	0.49	0.27	0.11
MKT PV.	(0.08)		(0.65)	(0.28)	(0.19)	(0.40)	
SMB C.	0.01	-0.08		0.71	0.001		0.1395
SMB PV.	(0.65)	(0.75)		(0.12)	(0.10)		
SMB C.	0.006	-0.12		0.77	0.16	0.28	0.17
SMB PV.	(0.82)	(0.65)		(0.11)	(0.69)	(0.41)	
HML C.	0.02	0.11	0.16		-0.32		0.3029
HML PV.	(0.04)	(0.34)	(0.13)		(0.05)		
HML C.	0.02	0.13	0.17		-0.39	-0.15	0.3343

HML PV.	(0.03)	(0.28)	(0.11)		(0.03)	(0.36)	
UMD C.	0.01	0.16	0.001	-0.57			0.2267
UMD PV.	(0.39)	(0.29)	(0.10)	(0.05)			
UMD C.	0.02	0.18	0.05	-0.56		-0.40	0.4058
UMD PV.	(0.22)	(0.19)	(0.69)	(0.03)		(0.03)	
IMU C.	0.02	0.14	0.13	-0.31	-0.58		0.2603
IMU PV.	(0.29)	(0.40)	(0.41)	(0.36)	(0.03)		

C.: *Coefficient.* ***PV.:*** *P-Value.* ***R2:*** *R-squared.*
Source: *Research finding.*

Table 10. 5-factor regression and the IMU

Variable	α	MKT	SMB	HML	RMW	CMA	IMU	R2
MKT C.	-0.04		-0.04	0.38	0.21	-0.22		0.036
MKT PV.	(0.12)		(0.87)	(0.49)	(0.65)	(0.58)		
MKT C.	-0.04		-0.04	0.43	0.28	-0.22	0.12	0.043
MKT PV.	(0.12)		(0.87)	(0.46)	(0.59)	(0.59)	(0.72)	
SMB C.	0.01	-0.04		0.11	-0.64	0.57		0.242
SMB PV.	(0.63)	(0.87)		(0.84)	(0.13)	(0.57)		
SMB C.	0.01	-0.04		0.13	-0.67	0.23	0.04	0.242
SMB PV.	(0.65)	(0.87)		(0.83)	(0.19)	(0.58)	(0.91)	
HML C.	0.02	0.07	0.02		-0.55	0.21		0.499
HML PV.	(0.13)	(0.49)	(0.84)		(0.002)	(0.20)		
HML C.	0.02	0.07	0.02		-0.60	0.20	-0.14	0.529
HML PV.	(0.13)	(0.46)	(0.83)		(0.001)	(0.23)	(0.30)	
RMW C.	0.007	0.05	-0.17	-0.74		0.44		0.635
RMW PV.	(0.59)	(0.65)	(0.64)	(0.002)		(0.02)		
RMW C.	0.008	0.06	-0.14	-0.74		0.39	-0.23	0.685
RMW PV.	(0.54)	(0.59)	(0.19)	(0.001)		(0.03)	(0.11)	
CMA C.	-0.03	-0.07	0.08	0.39	0.61			0.30
CMA PV.	(0.06)	(0.58)	(0.57)	(0.20)	(0.02)			
CMA C.	-0.03	-0.08	0.08	0.40	0.61		0.008	0.30
CMA PV.	(0.07)	(0.59)	(0.58)	(0.23)	(0.03)		(0.97)	
IMU C.	0.007	0.06	0.02	-0.43	-0.59	0.01		0.204
IMU PV.	(0.72)	(0.72)	(0.91)	(0.30)	(0.11)	(0.97)		

C.: *Coefficient.* ***PV.:*** *P-Value.* ***R2:*** *R-squared.*
Source: *Research finding.*

Appendix B: Tables for Statistical significance of factors in the multi-factorial models for stock returns

Table 11. Test results of the 3-factor model

Variable	Coefficients	t-statistics	P-Value	Standard error
HML	0.4	2.813	0.005	0.142
SMB	0.516	6.952	0.000	0.074
MKT	0.713	9.163	0.000	0.078
α	-0.046	-5.758	0.000	0.008

α: *Regression constant.*
Source: *Research finding.*

Table 12. Test results of the 3-factor model plus the factor IMU

Variable	Coefficients	t-statistics	P-Value	Standard error
HML	0.388	2.768	0.006	0.14
SMB	0.45	6.072	0.000	0.074
MKT	0.688	8.954	0.000	0.077
IMU	0.514	5.326	0.000	0.096
α	-0.051	-6.476	0.000	0.008

α: *Regression constant.*
Source: *Research finding.*

Table 13. Test results of the 4-factor model

Variable	Coefficients	t-statistics	P-Value	Standard error
HML	0.127	0.818	0.414	0.156
SMB	0.516	7.012	0.000	0.074
MKT	0.791	9.954	0.000	0.079
UMD	-0.477	-4.114	0.000	0.116
α	-0.039	-4.911	0.000	0.008

α: Regression constant.
Source: Research finding.

Table 14. Test results of the 4-factor model plus the factor IMU

Variable	Coefficients	t-statistics	P-Value	Standard error
HML	0.255	1.613	0.107	0.158
SMB	0.462	6.218	0.000	0.074
MKT	0.732	9.092	0.000	0.08
UMD	-0.236	-1.798	0.073	0.131
IMU	0.418	3.808	0.000	0.11
α	-0.047	-5.722	0.000	0.008

α: *Regression constant.*
Source: *Research finding.*

Table 15. Test results of the 5-factor model

Variable	Coefficients	t-statistics	P-Value	Standard error
HML	-0.178	-0.972	0.331	0.183
SMB	0.381	4.884	0.000	0.078
MKT	0.741	9.546	0.000	0.078
RMW	-0.759	-4.804	0.000	0.158
CMA	0.142	1.056	0.291	0.135
α	-0.046	-5.362	0.000	0.009

α: *Regression constant.*
Source: *Research finding.*

Table 16. Test results of the 5-factor model plus the factor IMU

Variable	Coefficients	t-statistics	P-Value	Standard error
HML	-0.014	-0.076	0.939	0.187
SMB	0.374	4.818	0.000	0.078
MKT	0.717	9.257	0.000	0.077
RMW	-0.536	-3.176	0.002	0.169
CMA	0.138	1.029	0.304	0.134
IMU	0.378	3.58	0.0004	0.106
α	-0.049	-5.68	0.000	0.009

α: *Regression constant.*
Source: *Research finding.*