# The effect of asymmetric information risk on returns of stocks traded on the BM&FBOVESPA

**Leonardo Souza Siqueira**

Universidade Federal de Minas Gerais, Faculdade de Ciências Econômicas, Centro de Pós-Graduação e Pesquisas em Administração, Belo Horizonte, MG, Brazil

**Hudson Fernandes Amaral**

Universidade Federal de Minas Gerais, Faculdade de Ciências Econômicas, Centro de Pós-Graduação e Pesquisas em Administração, Belo Horizonte, MG, Brazil
Universidade de Lisboa, Instituto Superior de Economia e Gestão, Departamento de Gestão, Lisboa, Portugal

**Laíse Ferraz Correia**

Centro Federal de Educação Tecnológica de Minas Gerais, Programa de Pós-Graduação em Administração, Belo Horizonte, MG, Brazil

## ABSTRACT

This study sought to analyze information asymmetry in the Brazilian stock market and its relation with the returns required from portfolios through the metrics volume-synchronized probability of informed trading. To do this, the study used actual data from the transactions of 142 stocks on the Brazilian Securities, Commodities and Futures Exchange (BM&FBOVESPA), within the period from May 1, 2014, to May 31, 2016. The results point out a high flow toxicity level in the orders of these stocks. In segment analyses of the stock market listing, data suggest there is no clue that stocks from the theoretically more overt segments have a lower toxicity level of order flows. The justification for this finding lies on the negative correlation observed between the market value of stocks and the toxicity level of orders. To test the effect of asymmetric information risk on stock returns, a factor related to the toxicity level of orders was added to the three-, four-, and five-factor models. Through the GRS test, we observed that the combination of factors that optimize the explanation of returns of the portfolios created was the one taking advantage of the factors market, size, profitability, investment, and information risk. To test the robustness of these results, the Average F-test was used in data simulated by the bootstrap method, and similar estimates were obtained. It was observed that the factor related to the book-to-market index becomes redundant in the national scenario for the models tested. Also, it was found that the factor related to information risk works as a complement to the factor size and that its inclusion leads to an improved performance of the models, indicating a possible explanatory power of information risk on portfolio returns. Therefore, data suggest that information risk is priced in the Brazilian stock market.

**Keywords:** information risk, return on assets, PIN, VPIN, asset pricing.

**R. Cont. Fin.** – USP, São Paulo, v. 28, n. 75, p. 425-425, set./dez. 2017

425

# 1. INTRODUCTION

Due to the increasing number of very frequently traded stocks and the concomitant expansion of tick-by-tick databases, market microstructure research has become increasingly viable. Particularly, this makes it possible for the microstructure field to be no longer seen only as a means of studying short-term asset price behavior, thus it is associated with other areas of finance studies, such as asset pricing.

The market microstructure area addresses the process and the consequences of buying and selling stocks (O'Hara, 1995). The main difference of this area to the traditional approach of the pricing models stems from the microstructure focus when analyzing how specific transaction mechanisms affect stock price formation. Therefore, one aspect of the microstructure is studying information content provided by stock prices. The difference between the information that market makers have in a market is named as information asymmetry and this has been the subject of studies since at least the 1970s.

Fama (1970) was a pioneer in the study of the role of the information set owned by shareholders by establishing the market efficiency difference in three ways: weak, semi-strong, and strong, depending on how the asset price reflects information about it. The models devised by Kyle (1985) and Glosten and Milgrom (1985) emerged in this context, and they propose one of the early market microstructure models by considering the effects of inside trading on bid and ask prices, from the market maker perspective.

Starting from Easley and O'Hara (1987), Easley, Kiefer, O'Hara and Paperman (1996) sought to quantify the information asymmetry observed in stock prices. Since then, several studies, mainly conducted by the above-mentioned authors (Easley, Engle, O'Hara, & Wu, 2008; Easley, López de Prado, & O'Hara, 2011) have sought to refine and develop a way of information asymmetry observed in a stock market, resulting first in the probability of informed trading (PIN) and later in the volume-synchronized probability of informed trading (VPIN). The VPIN seeks to directly measure the toxicity level of a stock's order flow. The term toxicity refers to the expected loss of a market maker by being in the same environment as a better informed agent, i.e. the more toxic an order flow is, the greater the probability that an individual with privileged issues purchase or sale at the same time as other investors provide liquidity, which results in an imbalance of orders.

Regarding the Brazilian stock market as riskier than developed country markets (Martins & Paulo, 2013) and taking into account that emerging countries are fertile ground for transactions driven by insiders (Duarte & Young, 2009), several scholars have proposed to study information asymmetry in the national market, both through the PIN (Barbedo, Silva, & Leal, 2009; Martins & Paulo, 2013, 2014) and alternative models (Iquiapaza, Lamounier, & Amaral, 2008; Albanez & Valle, 2009; Albanez, Lima, Lopes, & Valle, 2010). The empirical evidence of such research converges to the same result: there is a high probability that the inside trading practice is observed in the Brazilian stock market.

Information imbalance about assets traded in a financial market poses a risk to investors, who might, therefore, ask for a premium to trade those assets they perceive as riskier in terms of information level. Thus, the information risk of an asset may be one of the factors priced by market makers. The calculation of risk-adjusted rate of return on stocks generates controversy in the literature, and several models that propose to measure it have emerged, going through the studies by Sharpe (1964), Merton (1973), Jagannathan and Wang (1996), Ross (1976), Fama and French (1993, 2015), among others. A difficulty in its measurement lies on determining the explanatory factors that constitute the model, and the market factor is derived from the capital asset pricing model (CAPM), the most frequently used for asset pricing in the financial market (Fortunato, Motta, & Russo, 2010).

Despite their extensive use in practice and in finance research, Easley, Hvidkjaer and O'Hara (2005) point out inconsistency in the use of models such as the CAPM to study information asymmetry pricing by investors. This is due to the fact that the PIN and VPIN models derive from a scenario where participants have different information access levels, therefore, this factor might violate the aforementioned assumption, indicating the need to use other models.

Thus, this research followed the steps of Easley et al. (2005) and Mohanram and Rajgopal (2009), who study the influence of information risk on the required return on stocks by means of the introduction of a factor related to the PIN in the model by Fama and French (1993). This study goes a step further by estimating, through actual data, the effect of the VPIN on the required return on stocks traded in the Brazilian stock market, using the 5-factor model proposed by Fama and French (2015). Considering the need for further studies on the existence of insiders and their influence on the Brazilian stock

market, this research aimed to verify whether the stock order flow level, quantified by the VPIN, is a systematic risk factor priced by investors in shares traded on the Brazilian Securities, Commodities and Futures Exchange (BM&FBOVESPA) between May 1, 2014 and May 31, 2016.

This article is divided into 5 sections. The second presents a general literature review that grounded this empirical study; the third describes the methodologies used; and the fourth analyzes the results of estimated models. The last section resumes the objectives and reports our final remarks.

# 2. THEORETICAL REFERENCE

## 2.1 Information Asymmetry

The concept of information asymmetry is defined by Lambert, Leuz and Verrecchia (2011) as a result of the fact that a group of investors do not have access to information that is available to other participants. The use of such information for the purchase and sale of stocks in the financial market is called inside trading.

According to Leland (1992), numerous markets are characterized by imbalance of information between buyers and sellers. This phenomenon is even more pronounced in financial markets, especially in the relationship between borrowers and creditors. Taking into account the market efficiency hypothesis (MEH), proposed by Fama (1970), Leland (1992) points out the arguments for and against the inside trading practice. On the one hand, to the extent that stock prices reflect all available information (public and private), the insider's action causes new information to be incorporated into asset prices. On the other hand, potential investors are averse to entering this market because they consider it unfair. Thus, investments and asset prices and liquidity are smaller, affecting those investors who operate in the market without privileged information.

Having in mind the divergent scholars' conclusions about the effect of information asymmetry, it is interesting to analyze on a statistical basis the pertinence of its effect, which requires a means of quantifying this phenomenon.

## 2.2 Probability of Informed Trading

The PIN was introduced by Easley et al. (1996). First, it is assumed that asset purchase and sale operations occur on the basis of information held by investors. The authors claim that the model is based on the fact that, throughout the day, an informative event occurs randomly and is independently distributed, taking place with a probability $\alpha$; $\delta$ represents the probability that the information is bad news and $(1 - \delta)$, good news. Informed traders operate at a $\mu$ trading rate. In turn, uninformed traders operate in the

market with a $\varepsilon_b$ arrival rate for purchases and $\varepsilon_s$ for sales.

Easley et al. (2005) argue that the total physical volumes of buying and selling negotiations are sufficient for estimating the PIN. Thus, the model parameters, i.e. the vector $\theta = (\alpha, \mu, \varepsilon_b, \varepsilon_s, \delta)$, can be estimated through maximization of a maximum likelihood function.

As this is a probabilistic model that involves the occurrence of a set of various related events, the probability of negotiation with private information (PIN) follows the formulation (1):

$$PIN = \frac{\alpha\mu}{\alpha\mu + \varepsilon_b + \varepsilon_s} \qquad \boxed{1}$$

Since this is a model that uses intraday data and directly evaluates the probability of insiders' action, the PIN has been widely used in the financial literature and it has been empirically tested in the U.S. (Easley et al., 2005; Easley, Hvidkjaer, & O'Hara, 2002); Spanish (Abad & Rubia, 2005); Brazilian (Barbedo et al., 2009; Martins & Paulo, 2013, 2014; Agudelo, Giraldo, & Villaraga, 2015); French (Aktas, Bodt, Declerck, & Van Oppens, 2007); South Korean (Hwang, Lee, Lim, & Park, 2013); Colombian, Argentinean, Chilean, Peruvian, and Mexican markets (Agudelo et al., 2015), among others.

Duarte and Young (2009) point out that the PIN condenses the reasons that lead investors to launch transaction orders in only two: privileged information or search for liquidity. In a more recent study, Duarte, Hu and Young (2015) suggest that the PIN cannot capture insider information due to its operation mechanics.

In the Brazilian market, Barbedo et al. (2009) studied the relation between the PIN and the BM&FBOVESPA corporate governance levels. In turn, Martins and Paulo (2013, 2014) applied the PIN to the Brazilian market in the periods 2010 and 2011, seeking to relate the result found with the corporate governance levels and the companies' economic and financial characteristics, such as: risk, return, liquidity, cost of capital, and firm size, among others. The authors find a 25% average probability of

**R. Cont. Fin.** – USP, São Paulo, v. 28, n. 75, p. 425-444, set./dez. 2017

427

privileged transactions for companies within the analyzed period, a value higher than that found by Barbedo et al. (2009), i.e. 12.5%.

Due to many criticisms of the PIN, Easley et al. (2011) developed a new model for estimating the toxicity of transaction order flows, called VPIN.

## 2.3 Volume-Synchronized Probability of Informed Trading

Despite the extensive empirical application of the PIN and its relevance in the finance field – see a review by Mohanram and Rajgopal (2009) –, its calculation poses issues such as non-convergence of the maximum likelihood function for days in which the number of

orders is high. Thus, Easley, López de Prado and O'Hara (2012) proposed the volume-synchronized probability of informed trading (VPIN) based on Easley et al. (2008). This metrics, in addition to solving the issue mentioned above, seeks to directly quantify the toxicity level of order flows with no need for parameter estimation by means of maximum likelihood functions.

Easley et al. (2008) show that the expected value for the sum of purchase and sale volumes is equal to the total amount traded, represented by the denominator of equation (1). At the same time, the difference between the buying and selling volume may be the approximate value of informed traders' rate multiplied by the probability of an information event occurring, i.e. the numerator of (1). These relations are represented by equations (2) and (3):

$$E[V_t^B + V_t^S] = \alpha\mu + 2\varepsilon = nV \qquad \boxed{2}$$

$$E[V_t^B - V_t^S] \approx \alpha\mu \qquad \boxed{3}$$

The VPIN idea lies on dividing the day into equal volume buckets, treating each one equivalent to an information arrival period. Thus, $V_\tau^B + V_\tau^S$ is constant

and equal to V for every τ. Then, transaction imbalance is approximated by the average value calculated on volume buckets. So, the VPIN may be calculated by (4):

$$\text{VPIN} = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} \approx \frac{\sum_{\tau=1}^{n} |V_\tau^S - V_\tau^B|}{nV} \qquad \boxed{4}$$

## 2.4 Asset Pricing Models

Several asset pricing models have been proposed in the literature over the years, and the CAPM, proposed by Sharpe (1964), is one of the first and more impressive of them. The CAPM assumes that the expected return on assets is a linear function of its beta, multiplied by the market risk premium plus a risk-free asset return.

Based on the CAPM, other studies have been developed to increase the original model's robustness, such as the intertemporal CAPM (ICAPM), proposed by Merton (1973), the consumption-based CAPM, proposed by Breeden (1979), and the conditional CAPM (C- CAPM), proposed by Jagannathan and Wang (1996). Tests in international markets show that the C-CAPM has not been able to explain anomalies in asset returns (Lewellen & Nagel, 2006). In the Brazilian market, Tambosi Filho, Garcia, Imoniana and Moreiras (2010) and Flister, Bressan and Amaral (2011) do not find evidence contrary to the

C-CAPM, however, they recommend caution to use it, mainly due to national market immaturity. Machado, Bortoluzzo, Sanvicente and Martins (2013) analyze the ICAPM application in Brazil and verify that the results are favorable to the model within the period between 2003 and 2011.

Therefore, evidence was inconclusive regarding the most appropriate model for pricing shares. With the evolution of asset pricing studies, other risk sources were incorporated into the explanation of stock returns. Fama and French (1993), e.g. find there are at least three factors affecting returns on the assets analyzed. They are the so-called *small minus big* (SMB), *high minus low* (HML), and the market factor, which resulted in the 3-factor model proposed by Fama and French (1993).

Despite the apparent success of the 3-factor model when compared to the CAPM, Fama and French (1996) notice it is not able to explain returns on all assets and portfolios. Thus, Carhart (1997) suggests the addition

of a fourth factor, named as moment factor *up minus down* (UMD).

Due to evidence that emerged in the literature over the years that the 3 and 4-factor models are not able to explain variation in average returns related to profitability and investment, Fama and French (2015) revisit their previous model, adding 2 factors to it. The first of them, called *robust minus weak* (RMW), is obtained by the difference between returns on stock portfolios with high and low profitability. In turn, the investment-related factor, named as *conservative minus aggressive* (CMA), is the difference between return on low and high investment stock portfolios. The model in its complete form is represented by (5):

$$R_i - R_f = \alpha_i + b_i MKT + s_i SMB + h_i HML + r_i RMW + c_i CMA + \varepsilon_i \qquad \boxed{5}$$

Fama and French (2015) provide several contributions in relation to their previous model, e.g. the possibility of factor creation by means of combinations different from those used by Fama and French (1993, 1996). In order to corroborate the decision to use the model (5), Fama and French (2015) show that the value of the statistics proposed by Gibbons, Ross and Shanken (1989) – GRS – is lower for the 5-factor model than for the 3-factor model.

Fama and French (2016a) state that the anomalies related to the CAPM application decrease when the 5-factor model is applied. In addition, the latter has managed to solve problems related to the 3-factor model, such as those related to repurchase of stocks. In tests conducted in international markets, Fama and French (2016b) attest the 5-factor model superiority in relation to the others, but considering the failures related to the explanation of returns on small capitalization stocks. Other studies in markets such as the Australian (Chiah, Chai, Zhong, & Li, 2016), Japanese (Kubota & Takehara, 2017), Chinese (Lin, 2017), English (Nichol & Dowling, 2014), in addition to a study that gathers various European national markets (Zaremba & Czapkiewicz, 2016), show the superiority of this model in relation to the others.

Fama and French (2017) test variations of the factors proposed by Fama and French (2015) and show that the choice of factors are responses to empirical issues of the CAPM and C-CAPM. Therefore, the choice of factors is related to the discovery of patterns in asset returns and, to the extent that such patterns change over time, new factors may be added to the models.

Based on literature-based pricing models, several authors looked for a relation between the probability of privileged trading in a market and the required return on shares. The results found by the authors have differences, and in some studies a positive relation between the PIN and the required return was verified and, in others, there was no relation between them.

## 2.5 PIN and the Asset Pricing Models

Among the studies that sought to incorporate an information risk factor into asset pricing models, we highlight Easley et al. (2002, 2005), Mohanram and Rajgopal (2009), and Hwang et al. (2013), where the scholars proposed to empirically analyze the NIP influence on the required return on shares traded in the U.S. stock market. While Easley et al. (2002, 2005) and Hwang et al. (2013) claim that information risk is priced by investors through a systematic risk factor related to it, Mohanram and Rajgopal (2009) found contradictory results to this assertion.

**R. Cont. Fin.** – USP, São Paulo, v. 28, n. 75, p. 425-444, set./dez. 2017

429

Easley et al. (2005) create a PIN factor and add it to the 3-factor model proposed by de Fama and French (1993) and the 4-factor model proposed by Carhart (1997). The results show a statistically insignificant intercept for 8 out of the 10 portfolios when the PIN factor was added to the regression. By means of actual data, Hwang et al. (2013) calculate the PIN and regress it with the expected return represented by 4 different estimates of implicit cost of equity. The authors arrive at empirical results that support the hypothesis that there is a relation between information risk and expected returns, as reported by Easley et al. (2005).

Mohanram and Rajgopal (2009) replicating the studies conducted by Easley et al. (2002) conclude that returns on the PIN factor are negatively correlated with returns on stocks with a high PIN. Also, the PIN factor did not show a significant coefficient in the test with the 3- and 4-factor models proposed by Fama and French (1993) and Carhart (1997).

In addition to these works, Brennan, Huh and Subrahmanyam (2015) find evidence of information asymmetry pricing in the U.S. market through decomposition of the PIN into two factors. Borochin and Rush (2016), using the VPIN for the creation of a pricing factor, find favorable results to the hypothesis that there is an effect related to the information risk priced by market makers. Lai, Ng and Zhang (2014) show evidence contrary to the explanatory power of the PIN factor when analyzing stocks from 47 countries. Like Duarte and Young (2009), these authors conclude that the PIN may be more related to change in the demand for liquidity of stocks than information content.

The results of these studies point out an even deeper need to analyze the relation between information risk and the required return on shares. In Brazil, e.g. Martins and Paulo (2014) found a positive relation between the PIN and cost of capital and return on the shares. However, these authors did not resort to actual data of the transaction orders, hence, there is a possibility of problems related to the classification of orders.

## 3. METHODOLOGY

### 3.1 Population and Sample

The object of this study consisted of all the stocks traded on the BM&FBOVESPA. For calculating the VPIN, the sample was restricted to those stocks that had at least one transaction per day between May 1, 2014, and May 31, 2016, period in which information was available on the BM&FBOVESPA market data. Thus, the number of assets available for calculating this variable was 142 shares (common and preferred shares). For the formation of factors proposed by Fama and French (1993, 2015) and Carhart (1997) the sample available was 349 shares.

In the sample, both the preferred and common shares from the same company were analyzed, since empirical evidence in Brazil indicates they carry different information contents. Martins and Paulo (2013) found lower average PIN values for common shares in relation to preferred shares, even after considering the shares' liquidity level. In this study, however, no emphasis was assigned to the difference of VPIN for the different classes of shares.

### 3.2 Data Collection and Processing

The main limitation of studies that have proposed to apply the PIN or, in this case, the VPIN, is related to misclassification of purchase and sales orders. Aiming to circumvent the issue, this research resorts to actual data on the volume transacted in the Brazilian market. There is no record, in the Brazilian literature, of any research that uses the BM&FBOVESPA market data for calculating the VPIN, which evidences the originality of this study in the field of finance market microstructure.

The main reason for infrequent use of this directory may lie on data processing difficulty. Data are available by means of text files, containing a lot of information, such as: price of the deal, amount traded, time schedule, offer condition, code of the brokers involved, order type indicator, purchase or sale, among other data. Thus, dealing with these files requires long hours and hard work, in addition to the need for greater computational power for the separation and filtering of information relevant to the application of models.

Data collected through the BM&FBOVESPA market data were processed exclusively in routines developed on the statistical software *R*. The various other information, such as excessive stock returns, size, book-to-market, profitability, assets, non-current liabilities, and other data required for the application of models were collected from the Bloomberg˚ and Quantum Axis˚ databases.

## 3.3 Research Hypotheses

To achieve the objectives of this research, some hypotheses were tested in relation to the variables studied; they are:

**Hypothesis 1:** The smaller the company size, the higher the toxicity level of order flows.

According to Easley et al. (1996), shares from big companies have greater coverage of analysts and also greater attention of investors. Thus, the probability of having privileged transactions is, in theory, lower for these shares, resulting in lower VPIN than that found for shares from small companies, as also verified by Abad and Yagüe (2012) and Wei, Gerace and Frino (2013).

**Hypothesis 2:** The BM&FBOVESPA listing segments have different VPIN values.

It was expected that the companies constituting the

various segments of the BM&FBOVESPA have lower VPIN, as verified by Barbedo et al. (2009) and Martins and Paulo (2013). Thus, the hypothesis established was that the VPIN for the NM segment is the lowest, followed by N2, N1, and, finally, the traditional.

**Hypothesis 3:** A factor related to the VPIN helps explaining portfolio returns.

It was expected that the addition of a VPIN factor in the 3- and 5-factor models proposed by Fama and French (1993, 2015) and the 4-factor model proposed by Carhart (1997) had a reduced general intercept in the portfolios analyzed by the GRS and Average F-test proposed by Hwang and Satchell (2014).

## 3.4 Factor Models and the VPIN

Table 1 displays the specifications of models estimated in this study, which were based on Easley et al. (2005) and Mohanram and Rajgopal (2009).

**Table 1** *Models tested and their respective empirical bases*

| Model | Model specification | Empirical basis |
|---|---|---|
| Three Factors | $R_i - R_f = \alpha_i + \beta_i MKT + s_i SMB + h_i HML$ | Fama and French (1993) |
| Three Factors + IMU | $R_i - R_f = \alpha_i + \beta_i MKT + s_i SMB + h_i HML + p_i IMU_i$ | Fama and French (1993) |
| Four Factors | $R_i - R_f = \alpha_i + \beta_i MKT + s_i SMB + h_i HML + u_i UMD$ | Carhart (1997) |
| Four Factors + IMU | $R_i - R_f = \alpha_i + \beta_i MKT + s_i SMB + h_i HML + u_i UMD + p_i IMU_i$ | Carhart (1997) |
| Five Factors | $R_i - R_f = \alpha_i + \beta_i MKT + s_i SMB + h_i HML + r_i RMW_i + c_i CMA$ | Fama and French (2015) |
| Five Factors + IMU | $R_i - R_f = \alpha_i + \beta_i MKT + s_i SMB + h_i HML + r_i RMW_i + c_i CMA + p_i IMU_i$ | Fama and French (2015) |

**Source:** *Prepared by the authors.*

## 3.5 Dependent Variables

The dependent variable consists of the average daily excessive return on the stock portfolios in relation to the CDI  [$R_i - R_f$], formed according to the procedure adopted by Fama and French (2015). For creating the portfolios, the main variable 'size' was retained and the

second component was exchanged between the other variables: book-to-market, profitability, investment, and VPIN. Following Sanvicente and Bellato (2004), portfolios with less than 6 stocks were excluded from the tests, because they were not sufficiently diversified. Table 2 summarizes this procedure and illustrates information from the portfolios created.

**R. Cont. Fin.** – USP, São Paulo, v. 28, n. 75, p. 425-444, set./dez. 2017

431

**Table 2** *Information on portfolios created through the intersection of stocks divided into groups based on the variables: size, book-to-market, investment, profitability, and VPIN*

| Variables used in the division of stocks to create portfolios | Number of portfolios | Minimum of stocks | Maximum of stocks | Average of stocks |
|---|---|---|---|---|
| Size (3) / *Book-To-Market* (4) | 12 | 6 | 25 | 14 |
| Size (3) / Investment (4) | 12 | 7 | 23 | 14 |
| Size (2) / Investment (3) / Profitability (2) | 12 | 7 | 18 | 13 |
| Size (3) / VPIN (4) | 10 | 6 | 29 | 11 |
| Size (4) / Profitability (4) | 16 | 6 | 15 | 10 |
| Total | 62 | 6 | 29 | 12 |

*The column 'Variables' indicates which variables are used to build the portfolios. The values in parentheses refer to the breakpoints used in the division of stocks, e.g. (3) indicates that the stocks were divided into 3 groups. The intersections between the stocks were analyzed according to the values of variables, thus the portfolios were formed. Therefore, the number of portfolios consists in the multiplication of breakpoints for each variable. Attention is drawn to the fact that 2 portfolios were excluded from the combination Size and VPIN.*
**Source:** *Prepared by the authors.*

## 3.6 Independent Variables

For the creation of a factor related to information risk, first, the stocks were divided into 3 groups having their market values as a basis. At the same time, the stocks were divided into 2 groups: low and high VPIN. Finally, we calculated the weighed return on each intersection (Table 3).

**Table 3** *Portfolios formed having the variables size and VPIN as a basis to create the factor IMU*

| Portfolio | Initials | Description |
|---|---|---|
| Small and Low | SL | Intersection between stocks from the group *small* for the variable size and stocks with low VPIN value |
| Small and High | SH | Intersection between stocks from the group *small* for the variable size and stocks with high VPIN value |
| Medium and Low | ML | Intersection between stocks from the group *medium* for the variable size and stocks with low VPIN value |
| Medium and High | MH | Intersection between stocks from the group *medium* for the variable size and stocks with high VPIN value |
| Big and Low | BL | Intersection between stocks from the group *big* for the variable size and stocks with low VPIN value |
| Big and High | BH | Intersection between stocks from the group *big* for the variable size and stocks with high VPIN value |

**Source:** *Prepared by the authors.*

Thus, the factor IMU was obtained as shown in (6):

$$IMU = \frac{SH + MH + BH}{3} - \frac{SL + ML + BL}{3} \tag{6}$$

The reasons supporting the creation of the IMU lie on the relation between the probability of privileged negotiations and the return on shares. Easley et al. (2002) found a positive correlation between these 2 variables. According to these authors, stocks with higher PIN values have higher required return and, consequently, higher cost of capital. In this study, we observed that the correlation between VPIN and return was 0.0141 with $p = 0.0003$. Despite the low value, it is possible to have a premium for investment in stocks with greater toxicity of the order flows.

For creating the other factors, the procedures conducted by Fama and French (1993, 2015) and Carhart (1997) were followed. The factor SMB used in the 3- and 4-factor models proposed by Fama and French (1993, 2015) and Carhart (1997) was calculated by taking return on the intersection of stocks with low market value and low, medium, and high book-to-market and subtracting return on the intersection of stocks with high market value and low, medium, and high book-to-market. The factor HML was created through return on the intersection of stocks with high book-to-market and the groups low and high market value subtracted from return on the intersection of stocks with low book-to-market and low and high market value. The factor moment was obtained by taking return on the intersection of stocks with high

432

**R. Cont. Fin.** – USP, São Paulo, v. 28, n. 75, p. 425-444, set./dez. 2017

past return and low and high market value, subtracted from return on the intersection of stocks with low past return and low and high market value.

For the factor size used in the 5-factor models proposed by Fama and French (2015), we took return on the intersection of stocks with low market value and the other factors, subtracted from return on the intersection of stocks with high market value and the other factors. The factor HML was created through return on the intersection of stocks with high book-to-market and low and high market value subtracted from return on the intersection of stocks with low book-to-market and low and high market value.

For the profitability-related factor (RMW), we took return on the intersection of stocks with high profitability and low and high market value, subtracted from return on the intersection of stocks with low profitability and low and high market value. Finally, for the investment factor (CMA), return on the intersection of stocks with low investment and low and high market value was subtracted from return on the intersection of stocks with high investment and low and high market value.

## 3.7 Bootstrap Portfolio Simulation

Bootstrap simulation was used to analyze which combination of factors optimizes the explanation of returns on the portfolios. The application of this method consisted in selecting returns on the portfolios created by resampling with replacement, and the models were estimated under each of them, thus calculating the coefficients of each regression. After obtaining the regression intercepts, the Average F-test was applied. The application of this test was needed, because data simulation generates linear dependence between them, making impossible an inversion of the covariance matrix with regression residuals, needed to use the GRS test. For the moments in which no simulated data were used, the GRS test was applied following the steps proposed by Fama and French (1996, 2015).

## 4. ANALYSIS OF RESULTS

### 4.1 VPIN Results

The result for VPIN calculation is displayed in Table 4. A comparison of the results provided by Barbedo et al. (2009) and Martins and Paulo (2013, 2014) faces difficulties inherent to the procedures used in each investigation. Barbedo et al. (2009) and Martins and Paulo (2013, 2014) applied the PIN, in addition to using algorithms to classify the purchase and sale orders. Thus, a comparison of studies in other markets using the VPIN.

**Table 4** *Descriptive statistics of the VPIN for the entire sample between 05/01/2014 and 05/31/2016*

| Number of stocks | Minimum | Maximum | Average | Standard deviation |
|---|---|---|---|---|
| 142 | 0 | 1 | 0.4548 | 0.2218 |

**Source:** *Prepared by the authors.*

### 4.2 VPIN by the BM&FBOVESPA Listing Segment

Regarding the different BM&FBOVESPA listing segments, it was expected that companies in the New Market (NM) segment, which have, in theory, greater transparency than those in the segments Level 1 (L1), Level 2 (L2), and Traditional (Trad), had lower VPIN. To analyze this hypothesis, daily VPIN was calculated for each segment (Table 5).

**Table 5** *Daily VPIN descriptive statistics per BM&FBOVESPA listing segment*

| Descriptive statistics | Listing segments | | | |
|---|---|---|---|---|
| | L1 | L2 | NM | Trad |
| Number of stocks | 24 | 12 | 94 | 12 |
| Minimum | 0.3168 | 0.3669 | 0.4270 | 0.3466 |
| Maximum | 0.4310 | 0.6001 | 0.5039 | 0.5695 |
| Average | 0.3761 | 0.4870 | 0.4700 | 0.4597 |
| Standard deviation | 0.0191 | 0.0355 | 0.0124 | 0.0379 |

**Source:** *Prepared by the authors.*

**R. Cont. Fin.** – USP, São Paulo, v. 28, n. 75, p. 425-444, set./dez. 2017

433

We observe that the Level 2 VPIN was the highest among the BM&FBOVESPA listing segments, followed by the New Market, Traditional, and finally Level 1. It is worth noticing that the number of stocks in each segment is very different, and New Market is the segment with the largest number of companies. This fact strongly impacts the results by segment, as the analysis by market value explains.

In order to statistically verify the difference between the average VPIN values for the segments, the Student's *t*-test was used. The results point out rejection of the null hypothesis for equal values regarding the average VPIN values for all segments. It was expected that the values presented by the NM would be significantly lower than those for the other segments, especially the Traditional. The results found were, however, contrary to expectations. It was found that the segment L1 had the lowest VPIN for the sample analyzed and that the segment L2 had the highest average VPIN within the period concerned.

## 4.3 VPIN Analysis by Market Value of Stocks

One of the main results of studies that applied the PIN and, more recently, the VPIN, is that the probability of privileged trading is lower for stocks from big companies. For the VPIN, e.g. the disparity between buying and selling volumes is not so pronounced for larger capitalization firms, thus reducing their flow toxicity level. Abad and Yagüe (2012) were the first to verify this relation in stocks traded in the Spanish market. Wei et al. (2013) and Yildiz, Van Ness and Van Ness (2016) found evidence to support such a claim in studies in the Australian and U.S. markets, respectively. Having this in mind, we sought to verify the relation between the VPIN and companies' value in the Brazilian stock market.

To investigate this relation, stocks were divided into 3 groups named as *small*, *medium*, and *large*, having about 47 stocks each, related to their average daily market value. The first clue relating the VPIN to company size came from the correlation between these 2 variables. The result for correlation was -0.3080 and $p = 0$. Such a value was expected, given the constant empirical evidence of the negative relation between VPIN and size. In order to deepen the analysis of this relation, the descriptive statistics for each group was calculated (Table 6).

**Table 6** *Descriptive statistics of daily VPIN per group, according to stock market capitalization*

| Descriptive statistics | Groups | | |
|---|---|---|---|
| | Small | Medium | Large |
| Minimum | 0.5582 | 0.3544 | 0.2793 |
| Maximum | 0.7134 | 0.4565 | 0.3528 |
| Average | 0.6364 | 0.3989 | 0.3164 |
| Standard deviation | 0.0284 | 0.0164 | 0.0123 |

**Source:** *Prepared by the authors.*

Table 6 shows the relation verified in the studies that applied the VPIN to comparisons between the market capitalization of stocks. There is a significant difference between the VPIN for low capitalization stocks in the sample when compared to medium and high capitalization stocks. For comparison purposes, Table 7 depicts the results obtained in other markets where the VPIN was applied.

**Table 7** *Comparison between the VPIN calculated by size groups in different markets*

| Markets | Groups | | | References |
|---|---|---|---|---|
| | Small | Medium | Large | |
| Brazil | 0.6364 | 0.3989 | 0.3164 | This study |
| Spain | 0.5320 | 0.4020 | 0.2540 | Abad and Yagüe (2012) |
| Australia | 0.8459 | 0.7249 | 0.5002 | Wei et al. (2013) |
| USA | 0.6154 | 0.4001 | 0.2378 | Yildiz et al. (2016) |

**Source:** *Prepared by the authors.*

In general, the results obtained for medium and large companies in the Spanish and U.S. markets were close to those in the Brazilian market. We may highlight that the VPIN for stocks from the Brazilian companies shows a behavior similar to that of companies in the markets mentioned above, with negative correlation to company size. In general, the VPIN calculated for the domestic market, i.e. 0.4548, is not so different from that of the Spanish (0.3960) and U.S. (0.4178) markets.

Returning to the analysis of the VPIN characteristics regarding company size, Figure 1 shows the VPIN behavior for each group. We can observe the substantial difference in the toxicity level of stocks concerning their market values. The *small* group showed a daily VPIN always above 0.55, while the *medium* group was around 0.40, with slight peaks reaching 0.45. The *large* group remained more stable, with the lowest standard deviation among the 3, with a maximum of 0.35.
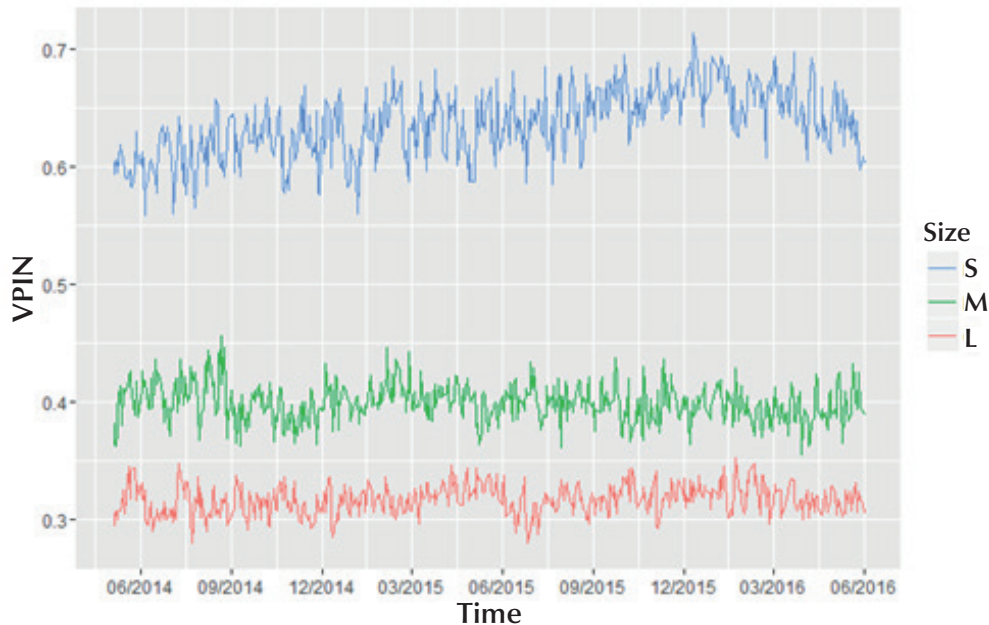


**Figure 1** *Evolution of daily VPIN for stock size groups.*
*S: small stocks group; M: medium stocks; L: large stocks in market value.*
**Source:** *Prepared by the authors.*

A substantial difference was found in the VPIN for stocks regarding the variable size, the explanation for VPIN behavior in the BM&FBOVESPA segments may be contained in the market values of stocks constituting each segment. Out of the 47 companies in the small group, 36 are in the New Market segment, which has 94 stocks in total, leading the average NM VPIN to increase substantially. Excluding the *small* group stocks, the NM VPIN would drop to 0.3418, a figure significantly lower than the current VPIN for the segment. For the L1 segment, among its 24 stocks, 1 is within the *small* group, 11 within the *medium* group, and 12 within the large company group, causing the VPIN to be taken down, and this might explain the fact that L1 shows the lowest VPIN among the segments. If the *small* and *medium* group stocks were excluded from the VPIN calculation in L1, their VPIN would drop from 0.3761 to 0.27972.

Regarding the segment L2, out of the 12 stocks that constitute it, 6 come from the *small* group, 3 from the *medium* group, and 3 from the *large* group. As the VPIN does not think of weight through company size and it assigns equal weight to companies, the fact that half of the segment consists of *small* group stocks might explain why L2 has the highest VPIN among the segments. In turn, for the Traditional segment, out of the 12 stocks that constitute it, 3 come from the *small* group, 3 from the *medium* group, and 6 from the *large* group. Among the stocks from the group with the highest market value there are the following assets: ABEV3, LAME3, LAME4, PETR3, PETR4, and VIVT4. Together, these stocks would have an average VPIN of 0.3030, far below the 0.4597 observed in the segment as a whole. The reason for such a difference lies on the VPIN for small stocks, which have an average value of 0.7835, leading the VPIN value for the Traditional segment to rise to 0.4597.

**R. Cont. Fin.** – USP, São Paulo, v. 28, n. 75, p. 425-444, set./dez. 2017

435

Therefore, the hypothesis that company size and its VPIN are negatively correlated was verified through the sample analyzed in the national stock market. Such evidence is in line with that expected and observed in the international literature.

## 4.4 Analysis of the 3-, 4- and 5-Factor Models and the Factor Based on the VPIN (IMU)

The first step to analyze the performance of models refers to the correlation between their factors (Table 8). We notice through panel (a), in Table 8, that the IMU showed a moderate and negative correlation to the market factor, which is close to that reported by Easley et al. (2005) and Mohanram and Rajgopal (2009). Regarding the factors HML and UMD, the IMU did not show a significant correlation, with $p = 0.1643$ and $0.9638$.

It is noticed that the correlation of greater weight

for the factor constructed by the VPIN is to the SMB. It is possible that return on stocks with higher VPIN are those of small companies, while stocks with lower VPIN represent the larger stocks, and the explanation for a strong correlation between the 2 factors lies on their construction.

When analyzing the correlations between the 5 factors proposed by Fama and French (2015) and the IMU – panel (b), in Table 8 –, it is observed that, even in different methodologies, the IMU has a slightly stronger correlation to the factor size. Another moderate and positive correlation arises between the IMU and the CMA. In general, the results do not resemble those presented by Fama and French (2015), except for the positive correlation between the factors CMA and HML. Again, the factors MKT and SMB show a moderate and negative correlation, while for Fama and French (2015) there is a correlation with the same magnitude, but positive.

**Table 8** *Correlation between the 3, 4 and 5 factors proposed by Fama and French (1993, 2015) and Carhart (1997) and the factor IMU*

| Panel (a): Correlation between the 3 factors proposed by Fama and French (1993), the factor moment proposed by Carhart (1997), and the factor IMU | | | | | |
|---|---|---|---|---|---|
| | **MKT** | **SMB** | **HML** | **UMD** | **IMU** |
| MKT | 1 | | | | |
| SMB | -0.4463 | 1 | | | |
| HML | 0.3806 | -0.1140 | 1 | | |
| UMD | -0.2801 | -0.0110 | -0.5750 | 1 | |
| IMU | -0.2892 | 0.5633 | 0.0614 | -0.0020 | 1 |
| Panel (b): Correlation between the 5 factors proposed by Fama and French (2015) and the factor IMU | | | | | |
| | **MKT** | **SMB** | **HML** | **RMW** | **CMA** | **IMU** |
| MKT | 1 | | | | | |
| SMB | -0.3334 | 1 | | | | |
| HML | 0.3806 | 0.0398 | 1 | | | |
| RMW | 0.0236 | -0.0104 | -0.4864 | 1 | | |
| CMA | 0.0885 | 0.3766 | 0.3489 | 0.0220 | 1 | |
| IMU | -0.2892 | 0.5650 | 0.0614 | -0.1106 | 0.2861 | 1 |

**Source:** *Prepared by the authors.*

### 4.4.1 Results of factor model regressions.

In order to present the results regarding the regressions estimated, there is a need to apply econometric tests that aim to test the statistical robustness of the models analyzed. Three tests were performed to verify, respectively, whether there is multicollinearity, whether the regression residues are autocorrelated, and whether the latter are heteroskedastic. The first test, variance inflation factor (VIF), tests whether the explanatory variables are correlated, something which might affect the estimation of their coefficients. Following Gujarati (2006), VIF values above 10 indicate multicollinearity. It

was verified that, in none of the 6 models, the result of VIF for the variables was high. Therefore, the results indicate there is no multicollinearity between the variables, and it is possible to include them in a regression model with no apparent loss in the coefficient estimates.

The Durbin-Watson and Breusch-Pagan tests were also applied. Regarding the first test, out of the 62 portfolios, only 1 rejected the null hypothesis of absence of autocorrelation between the residues, having presented a test statistics of 1.77, i.e. the correlation was positive, but it was not strong. Regarding the Breusch-Pagan test, only 6 portfolios rejected the null hypothesis of homoscedasticity.

Having these results in mind, it is concluded that the OLS estimators are sufficient for a concise estimation of the models' coefficients.

Table 9 shows the results for the GRS test applied to the portfolio sets. It is verified that, for the portfolios constructed having size and book-to-market as a basis, adding the factor IMU to the 3-, 4-, and 5-factor models leads to improvement in their performance. For the portfolios constituted by company size and investment, adding the IMU does not entail significant differences in relation to the traditional models. A detail to be highlighted is that the 5-factor model showed worse performance than the 3- and 4-factor models for this set of portfolios.

What can be verified is that no model has performed better regardless of the set analyzed. Fama and French

(2015) show that the 5-factor model has a significant improvement over the 3-factor model for the 7 portfolio sets analyzed by them. In this study, the 5-factor model was better in those sets created through the variables size and VPIN and size and profitability. Portfolio formation based on profitability presented the best explanation level considering the models as a whole. When formed by book-to-market, investment, or VPIN, the models did not perform satisfactorily.

Through the results presented by the GRS test, there are clues that an information-related factor, when added to the traditional factor models, is adequate to explain returns. In order to analyze this hypothesis more deeply, we regress the factors based on Mohanram and Rajgopal (2009) and Fama and French (2015).

**Table 9** *GRS test for the 3-, 4-, and 5-factor models plus the factor IMU*

| Combination of factors | Portfolios tested | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 12 portfolios: Size and BM. | | 12 portfolios: Size and Inv. | | 12 portfolios: Size, Inv. and Profit. | | 10 portfolios: Size and VPIN | | 16 portfolios: Size and Profit. | |
| | GRS | *P* value | GRS | *P* value | GRS | *P* value | GRS | *P* value | GRS | *P* value |
| 3 Factors | 1.7726 | 0.0498 | **1.6089** | **0.0853** | 1.0109 | 0.4372 | 1.8052 | 0.0570 | 0.8763 | 0.5970 |
| 3 Factors + IMU | **1.7336** | **0.0568** | 1.6116 | 0.0846 | **0.9900** | **0.4571** | 1.7529 | 0.0666 | 0.8844 | 0.5876 |
| 4 Factors | 1.7895 | 0.0470 | 1.6112 | 0.0847 | 1.0108 | 0.4372 | 1.8050 | 0.0571 | 0.8762 | 0.5973 |
| 4 Factors + IMU | 1.7554 | 0.0528 | 1.6121 | 0.0845 | 0.9913 | 0.4558 | 1.7538 | 0.0664 | 0.8854 | 0.5864 |
| 5 Factors | 1.7754 | 0.0493 | 1.6559 | 0.0734 | 1.0258 | 0.4232 | 1.7715 | 0.0630 | **0.8589** | **0.6176** |
| 5 Factors + IMU | 1.7459 | 0.0545 | 1.6699 | 0.0701 | 1.0076 | 0.4403 | **1.7318** | **0.0708** | 0.8677 | 0.6072 |

*Size: Size; BM: Book-to-Market; Inv.: Investment; Profit.: Profitability.*
*The best combination of factors for each portfolio set is highlighted in bold.*
**Source:** *Prepared by the authors.*

### 4.4.2 Analysis of the relations between systematic risk factors.

This procedure, performed by Mohanram and Rajgopal (2009) and Fama and French (2015), aims to test whether the regression intercepts are statistically different from 0. An intercept equal to 0 would mean that the factor is not priced and that its predictive power is already incorporated to the existing factors.

Table 10 shows the estimates for the 3-factor regressions proposed by Fama and French (1993) and the factor IMU. The factor SMB presents an intercept statistically different from 0, i.e. the other factors do not incorporate it to returns. When the IMU is added, there is a substantial increase in the regression $R^2$. This is probably due to the strong positive correlation of 0.56 between the 2 elements. However, despite this increased explanatory power of

regression, the factor SMB is still not captured by the others.

Analyzing the factor HML, this is the one with the highest $p$ value for its intercepts. The factors MKT and IMU are those that show greater explanation power on the HML. Fama and French (2015) show that the HML becomes redundant for the 5-factor model. The results in Table 10 suggest that, even for the 3-factor model, the HML does not show to be a relevant factor, given its incorporation, particulary, by the market factor. Evidence using the GRS test shows that excluding HML from the 3-factor model, leaving only the factor market and the SMB, increases the model's explanatory power, as evidenced in the next section. For the factor IMU, the null hypothesis regarding its intercept is not rejected. This seems to be incorporated by the other factors, especially the SMB.

**R. Cont. Fin.** – USP, São Paulo, v. 28, n. 75, p. 425-444, set./dez. 2017

437

**Table 10** *Three-factor regression proposed by Fama and French (1993) and the IMU*

| Variable regressed | | α | Independent variables | | | | R² |
|---|---|---|---|---|---|---|---|
| | | | MKT | SMB | HML | IMU | |
| MKT | Coeff. | -0.0008 | | -0.5627 | 0.3885 | | 0.3067 |
| | P value | (0.148) | | (0) | (0) | | |
| | Coeff. | -0.0008 | | -0.4674 | 0.4062 | -0.2183 | 0.3149 |
| | P value | (0.188) | | (0) | (0) | (0) | |
| SMB | Coeff. | -0.001 | -0.3418 | | 0.0515 | | 0.1998 |
| | P value | (0.003) | (0) | | (0.126) | | |
| | Coeff. | -0.001 | -0.2146 | | -0.026 | 0.6349 | 0.4023 |
| | P value | (0.002) | (0) | | (0.412) | (0) | |
| HML | Coeff. | -0.004 | 0.3539 | 0.0827 | | | 0.1454 |
| | P value | (0.427) | (0) | (0.127) | | | |
| | Coeff. | -0.005 | 0.3623 | -0.0505 | | 0.3253 | 0.1732 |
| | P value | (0.334) | (0) | (0.413) | | (0) | |
| IMU | Coeff. | 0 | -0.0633 | 0.4004 | 0.105 | | 0.3387 |
| | P value | (0.314) | (0) | (0) | (0) | | |

Coeff.: Coefficient.
**Source:** *Prepared by the authors.*

In the analysis of 4-factor regressions proposed by Carhart (1997), depicted in Table 11, it is verified that the factor UMD had the highest *p* values for its intercepts, i.e. 0.869 and 0.95, the latter refers to adding the factor IMU. This means there is strong evidence that the other factors completely capture the factor UMD.

**Table 11** *Four-factor regression proposed by Carhart (1997) and the IMU*

| Variable regressed | | α | Independent variables | | | | | R² |
|---|---|---|---|---|---|---|---|---|
| | | | MKT | SMB | HML | UMD | IMU | |
| MKT | Coeff. | 0 | | -0.5776 | 0.2938 | -0.2133 | | 0.3183 |
| | P value | (0.159) | | (0) | (0) | (0) | | |
| | Coeff. | 0 | | -0.4928 | 0.3177 | -0.2108 | -0.1913 | 0.3242 |
| | P value | (0.195) | | (0) | (0) | (0.004) | (0.019) | |
| SMB | Coeff. | -0.001 | -0.35 | | -0.0181 | -0.1906 | | 0.215 |
| | P value | (0.004) | (0) | | (0.666) | (0) | | |
| | Coeff. | -0.001 | -0.2227 | | -0.1033 | -0.2002 | 0.6376 | 0.4197 |
| | P value | (0.003) | (0) | | (0.005) | (0) | (0) | |
| HML | Coeff. | -0.0002 | 0.1976 | -0.0201 | | -0.7281 | | 0.3795 |
| | P value | (0.556) | (0) | (0.667) | | (0) | | |
| | Coeff. | -0.0003 | 0.2064 | -0.1485 | | -0.7251 | -0.3144 | 0.4059 |
| | P value | (0.433) | (0) | (0) | | (0) | (0) | |
| UMD | Coeff. | 0 | -0.0807 | -0.1099 | -0.3781 | | | 0.3451 |
| | P value | (0.869) | (0) | (0) | (0) | | | |
| | Coeff. | 0 | -0.0736 | -0.1547 | -0.39 | | 0.1119 | 0.3507 |
| | P value | (0.950) | (0.004) | (0) | (0) | | (0.02) | |
| IMU | Coeff. | 0 | -0.0557 | 0.4111 | 0.141 | 0.0933 | | 0.3443 |
| | P value | (0.320) | (0) | (0) | (0) | (0.020) | | |

Coeff.: Coefficient.
**Source:** *Prepared by the authors.*

Finally, concerning the two-factor regressions added by Fama and French (2015), depicted in Table 12, the HML is again captured by the other factors. The factors RMW and CMA showed intercepts close to 0 and they have a high *p* value.

Therefore, through the regression estimates of factors displayed in tables 10 to 12, it can verify that the 2 main factors responsible for explaining return on the portfolios analyzed seem to be the MKT and the SMB, and the others are incorporated as other factors are added to regressions.

In order to verify which combination of factors results in the best model, the procedure established by Fama and French (2015) is adopted, i.e. analysis of GRS test results for various model arrangements.

**Table 12** *Five-factor regression proposed by Fama and French (2015) and the IMU*

| Variable regressed | | α | Independent variables | | | | | | R² |
|---|---|---|---|---|---|---|---|---|---|
| | | | MKT | SMB | HML | RMW | CMA | IMU | |
| MKT | Coeff. | 0.002 | | -5294 | 0.5933 | 0.4237 | 0.0687 | | 0.3213 |
| | *P* value | (0.261) | | (0) | (0) | (0) | (0.324) | | |
| | Coeff. | 0.001 | | -0.4199 | 0.5817 | 0.4188 | 0.0958 | -0.2652 | 0.334 |
| | *P* value | (0.326) | | (0) | (0) | (0) | (0.168) | (0) | |
| SMB | Coeff. | -0.001 | -0.2722 | | 0.0517 | 0.0253 | 0.4315 | | 0.2739 |
| | *P* value | (0.004) | (0) | | (0.227) | (0.641) | (0) | | |
| | Coeff. | -0.001 | -0.1769 | | 0.04 | 0.0413 | 0.289 | 0.5318 | 0.4163 |
| | *P* value | (0.003) | (0) | | (0.296) | (0.145) | (0) | (0) | |
| HML | Coeff. | -0.0003 | 0.3266 | 0.0553 | | -0.7332 | 0.4262 | | 0.4945 |
| | *P* value | (0.463) | (0) | (0.227) | | (0) | (0) | | |
| | Coeff. | -0.0003 | 0.3269 | 0.0535 | | -0.7327 | 0.4257 | 0.0049 | 0.4935 |
| | *P* value | (0.462) | (0) | (0.296) | | (0) | (0) | (0.936) | |
| RMW | Coeff. | -0.0001 | 0.1551 | 0.0168 | -0.4554 | | 0.2105 | | 0.3295 |
| | *P* value | (0.593) | (0) | (0.641) | (0) | | (0) | | |
| | Coeff. | -0.0001 | 0.1446 | 0.0585 | -0.4502 | | 0.2203 | -0.1126 | 0.3354 |
| | *P* value | (0.662) | (0) | (0.145) | (0) | | (0) | (0.018) | |
| CMA | Coeff. | 0.0002 | 0.0277 | 0.3392 | 0.313 | 0.2489 | | | 0.2959 |
| | *P* value | (0.593) | (0.324) | (0) | (0) | (0) | | | |
| | Coeff. | 0.0001 | 0.0389 | 0.2786 | 0.3076 | 0.2591 | | 0.1483 | 0.3058 |
| | *P* value | (0.678) | (0.168) | (0) | (0) | (0) | | (0.004) | |
| IMU | Coeff. | 0.0002 | -0.078 | 0.3715 | 0.002 | -0.096 | 0.1075 | | 0.3471 |
| | *P* value | (0.364) | (0) | (0) | (0.935) | (0.018) | (0.004) | | |

*Coeff.: Coefficient.*
**Source:** *Prepared by the authors.*

**4.4.3 Results for the GRS test in various model combinations.**

To find the model that best explains portfolio returns, Fama and French (2015) test various variable arrangements. The same procedure was performed in order to verify which combination of factors produces the best model for the sample analyzed (Table 13).

**R. Cont. Fin.** – USP, São Paulo, v. 28, n. 75, p. 425-444, set./dez. 2017

439

**Table 13** *GRS test for 3-, 4-, and 5-factor model variations and the factor IMU*

| Combinations of factors | | Portfolios tested | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 12 portfolios: Size and BM. | | 12 portfolios: Size and Inv. | | 12 portfolios: Size, Inv. and Profit. | | 10 portfolios: Size and VPIN | | 16 portfolios: Size and Profit. | |
| | | GRS | P value | GRS | P value | GRS | P value | GRS | P value | GRS | P value |
| **3 Factors + IMU** | MKT, SMB, HML | 1.7726 | 0.0498 | 1.6089 | 0.0853 | 1.0109 | 0.4372 | 1.8052 | 0.0570 | 0.8763 | 0.5970 |
| | MKT, SMB | 1.7465 | 0.0543 | 1.5668 | 0.0975 | 0.9892 | 0.4578 | 1.7639 | 0.0645 | 0.8770 | 0.5962 |
| | MKT, SMB, HML, IMU | 1.7336 | 0.0568 | 1.6116 | 0.0846 | 0.9900 | 0.4571 | 1.7529 | 0.0666 | 0.8844 | 0.5876 |
| | MKT, SMB, IMU | 1.7096 | 0.0615 | **1.5576** | **0.1004** | **0.9664** | **0.4800** | **1.7148** | **0.0744** | 0.8841 | 0.5879 |
| **4 Factors + IMU** | MKT, SMB, HML, UMD | 1.7895 | 0.0470 | 1.6112 | 0.0847 | 1.0108 | 0.4372 | 1.8050 | 0.0571 | 0.8762 | 0.5973 |
| | MKT, SMB, HML, UMD, IMU | 1.7554 | 0.0528 | 1.6121 | 0.0845 | 0.9913 | 0.4558 | 1.7538 | 0.0664 | 0.8854 | 0.5864 |
| | MKT, SMB, UMD | 1.7379 | 0.0559 | 1.5881 | 0.0912 | 0.9897 | 0.4573 | 1.7688 | 0.0636 | 0.8742 | 0.5996 |
| | MKT, SMB, UMD, IMU | **1.7005** | **0.0634** | 1.5806 | 0.0934 | 0.9670 | 0.4794 | 1.7198 | 0.0734 | 0.8814 | 0.5911 |
| **5 Factors + IMU** | MKT, SMB, HML, CMA, RMW | 1.7754 | 0.0493 | 1.6559 | 0.0734 | 1.0258 | 0.4232 | 1.7715 | 0.0630 | **0.8589** | **0.6176** |
| | MKT, SMB, CMA, RMW | 1.7533 | 0.0531 | 1.6169 | 0.0832 | 1.0118 | 0.4364 | 1.7599 | 0.0652 | 0.8597 | 0.6166 |
| | MKT, SMB, HML, CMA, RMW, IMU | 1.7459 | 0.0545 | 1.6699 | 0.0701 | 1.0076 | 0.4403 | 1.7318 | 0.0708 | 0.8677 | 0.6072 |
| | MKT, SMB, CMA, RMW, IMU | 1.7237 | 0.0587 | 1.6284 | 0.0802 | 0.9924 | 0.4547 | 1.7198 | 0.0734 | 0.8684 | 0.6064 |
| | MKT, SMB, RMW | 1.7699 | 0.0502 | 1.6304 | 0.0797 | 1.0010 | 0.4465 | 1.7812 | 0.0613 | 0.8677 | 0.6072 |
| | MKT, SMB, RMW, IMU | 1.7337 | 0.0567 | 1.6373 | 0.0779 | 0.9813 | 0.4655 | 1.7312 | 0.0710 | 0.8739 | 0.5999 |
| | MKT, SMB, CMA | 1.7547 | 0.0529 | 1.6181 | 0.0829 | 1.0090 | 0.4390 | 1.7615 | 0.0649 | 0.8604 | 0.6158 |
| | MKT, SMB, CMA, IMU | 1.7230 | 0.0588 | 1.6259 | 0.0808 | 0.9851 | 0.4618 | 1.7191 | 0.0735 | 0.8684 | 0.6064 |

The column Models shows which factor combinations were used as independent variables in the explanation of portfolio returns (dependent variable in the regressions) evidenced in the other columns.

The best combination of factors for each portfolio set is highlighted in bold.

Size: Size; BM.: Book-to-Market; Inv.: Investment; Profit .: Profitability.

**Source:** *Prepared by the authors.*

Analyzing the set of portfolios formed by stock size and book-to-market, it is verified that, among the 16 factor arrangements analyzed, the one showing the best performance consisted in the factors MKT, SMB, UMD, and IMU. An issue arises when including the UMD improves the model performance, considering that it was the one with the lowest intercept and the highest *p* values in Table 11. The explanation lies on the relation between UMD and HML. Analyzing the UMD regression with the other factors, it is verified that the HML shows the highest absolute coefficient among the variables, hence indicating that this might be the factor that better captures the UMD return variations. In unreported results, regressions of this factor with the MKT, SMB, and later by adding the IMU, there is a substantial decrease in the *p* value of the intercept, and this suggests these factors cannot fully capture the factor UMD, leaving margin to play a role in explaining portfolio returns.

It is also verified that adding the IMU improves the performance of all models analyzed. Again, we should seek an explanation for this evidence, since, through factor regressions, it was found that the IMU was captured by the SMB and the HML. The SMB is the factor with the greatest explanatory power regarding the IMU. As the IMU aims to capture the informational part of stocks

and the relation to company size is a consequence of how the capital market deals with the companies' information content, it is possible that the IMU is able to explain a part of the portfolio return variations not captured by the factor SMB – the part related to information risk –, and this might explain the improved performance of models that include the factor IMU.

In conclusion, it was found that the best factor combinations were those that excluded the HML and included the IMU. The MKT, SMB, and IMU model showed the best performance for 3 out of the 5 portfolio sets. The MKT, SMB, UMD, and IMU model showed better performance for portfolios formed by size and book-to-market, but it also demonstrated good performance in tests with the other sets, except for the size and profitability combination. For the latter, the best combination consisted in the 5-factor model, followed by the 4-factor model (MKT, SMB, RMW, and CMA). The next section works with portfolio return simulations to check which of these factor combinations best explains returns in the sample analyzed.

### 4.4.4 Model analysis using the Bootstrap method.

This section highlights the best model among the following factor combinations: MKT, SMB, IMU; MKT,

SMB, UMD, IMU; MKT, SMB, HML, RMW, CMA; MKT, SMB, RMW, CMA. We chose to include the 5-factor models proposed by Fama and French (2015) and its restricted version without the factor HML, in order to verify, through simulations, the effect of HML exclusion on the model performance. To do this, the bootstrap method was used for simulating portfolio returns. To check the hypothesis that intercepts are not statistically different from 0, the Average F-test was used. Finally, in order to determine which of these models showed better overall performance, the Fisher's method was used to combine the models' $p$ values, as shown by (7):

$$-2 \sum_{i=1}^{n} \ln X_i \sim \chi^2(2n) \qquad \boxed{7}$$

where $\ln X_i$ is the natural logarithm of each $p$ value. The results for $p$ value combinations are shown in Table 14. For each portfolio set, one model outperformed the others.

Thus, $p$ value combinations for the 62 simulations – related to the 62 portfolios of the 5 sets formed.

Table 14 *Result of p value combinations by the Fisher's method*

| Combinations of factors | Portfolios tested | | | | | |
|---|---|---|---|---|---|---|
| | Size (3) and BM (4) | Size (3) and Inv. (4) | Size (2), Inv. (3) and Profit. (2) | Size (3) and VPIN (4) | Size (4) and Profit. (4) | General |
| MKT, SMB, IMU | $2.06 \times 10^{-7}$ | $4.22 \times 10^{-6}$ | **$1.38 \times 10^{-6}$** | $2.91 \times 10^{-5}$ | $1.92 \times 10^{-6}$ | $3.55 \times 10^{-24}$ |
| MKT, SMB, UMD, IMU | $2.11 \times 10^{-7}$ | $5.51 \times 10^{-6}$ | $1.16 \times 10^{-6}$ | $3.08 \times 10^{-5}$ | $2.09 \times 10^{-6}$ | $4.55 \times 10^{-24}$ |
| MKT, SMB, HML, RMW, CMA | $1.87 \times 10^{-7}$ | **$5.54 \times 10^{-6}$** | $1.13 \times 10^{-6}$ | $2.81 \times 10^{-5}$ | **$2.82 \times 10^{-6}$** | $4.06 \times 10^{-24}$ |
| MKT, SMB, RMW, CMA | **$2.29 \times 10^{-7}$** | $5.38 \times 10^{-6}$ | $1.11 \times 10^{-6}$ | **$3.21 \times 10^{-5}$** | $2.27 \times 10^{-6}$ | **$5.22 \times 10^{-24}$** |

The best factor combination for each portfolio simulation set is highlighted in bold.
Size: Size; BM.: Book-to-Market; Inv .: Investment; Profit.: Profitability.
**Source:** *Prepared by the authors.*

The model MKT, SMB, and IMU, although showing better results for the sets, as displayed in Table 14, did not support in face of the others when applied to simulated data. The inclusion of the factor moment, resulting in the model MKT, SMB, UMD, and IMU, however, managed to maintain a good performance in portfolio simulations, suggesting that the factor UMD can capture return variations in the general sample.

Finally, due to the evidence displayed in Table 13, that including the factor IMU instead of the factor HML, generally leads to improved model performance, a fifth model's performance, not reported in Table 14, was analyzed: MKT, SMB, RMW, CMA, and IMU. The result of $p$ value combinations in the simulations was $7.72 \times 10^{-24}$, i.e. superior to the performance of the model MKT, SMB, RMW, and CMA, i.e. $5.22 \times 10^{-24}$, which indicates that adding the factor related to information risk led to the improvement in the restricted model proposed by Fama and French (2015).

## 5. FINAL REMARKS

This study aimed to: (i) analyze the VPIN or the toxicity level of stocks in the Brazilian market; and (ii) verify, through factor models proposed by Fama and French (1993, 2015) and Carhart (1997), whether a systematic risk factor related to stocks' information content is priced by the BM&FBOVESPA investors.

An average VPIN value of 0.4548 was found, with a standard deviation of 0.2219 for the Brazilian market. In the analysis related to the stock listing segments, it was verified that the L1 segment showed lower VPIN, followed by the Traditional, NM, and L2. NM stocks were expected to have lower VPIN values, since the

**R. Cont. Fin.** – USP, São Paulo, v. 28, n. 75, p. 425-444, set./dez. 2017

441

BM&FBOVESPA segmentation objective is providing the investor with greater transparency, which would imply a lower probability of inside trading. The results suggest that the probability of inside trading in the segments is related to the number of companies and the characteristics of stocks that comprise them, especially their market value. Thus, the hypothesis stipulated in this research, that the theoretically more transparent segments of the BM&FBOVESPA might have lower probability of inside trading, could not be confirmed.

The hypothesis that there is a negative correlation between size and the VPIN of stocks was corroborated. The results of this study indicate that there is a correlation of -0.3080 between the market value and the companies' VPIN. The sample analyzed was divided into 3 groups related to the companies' market value: small, medium, and large. The average VPIN value for these groups was 0.6364, 0.3989, and 0.3164, respectively, indicating a clear decrease in the VPIN as company size increased.

The last hypothesis, regarding the role of a factor related to the information risk of stocks, was analyzed by constructing the factor IMU. This factor was added to the 3- and 5-factor models proposed by Fama and French (1993, 2015) and the 4-factor models proposed by Carhart (1997) and we used, as dependent variables in regressions, the returns on 62 portfolios constructed having size, book-to-market, profitability, investment, and the stock VPIN values as a basis. For all models, adding the factor IMU increased the predictive power of the SMB in the sample analyzed.

In general, the improved model performance by including the IMU was verified through the GRS test. In order to further investigate this assertion, regressions between the factors were performed. The results indicate that all factors, except the SMB and the MKT are captured by the others at some point. Fama and French (2015) notice that, in the 5-factor model context, the factor HML becomes redundant. In order to verify which of these factors help explaining portfolio returns, the GRS test was applied to various combinations of factors.

The results indicate that the following models showed better performance: MKT, SMB, and IMU; MKT, SMB, UMD, and IMU; MKT, SMB, HML, CMA, and RMW; and MKT, SMB, CMA, and RMW. In addition to corroborating the result found by Fama and French (2015), that the factor HML is redundant for the 5-factor model, this claim was extended to the 3- and 4-factor models, and it was found that, when present, the HML affects the models' performance.

In order to extend these conclusions, the bootstrap procedure of portfolio returns was carried out, being regressed through the models mentioned in the previous paragraph. Subsequently, the Average F-test was applied. The results for this test indicate that the model that best explains the simulated returns is MKT, SMB, RMW, and CMA. From the previously presented evidence that the factor IMU helps in the models' performance, we resorted to return simulations with the model MKT, SMB, RMW, CMA, and IMU. The result found was that the latter had a better performance than the other models, and this provides support for the central hypothesis of this study on information risk pricing in the Brazilian stock market.

Through the estimate results, it is understood that the factor IMU works as a complement to the factor SMB – the latter is key for the models' performance – related to the information risk of stocks. The explanation takes place by the way both of them are constructed, since small companies are strongly present in the informed group and the big companies constitute the uninformed group. If these two factors were proxies for each other, the VIF test would have a high value. In addition, it was verified that the correlation between the 2, although positive, is not enough for one factor to completely incorporate the other, a fact evidenced by the IMU regression with the other factors.

In conclusion, the factor related to information risk seems to play a significant role in explaining the return on portfolios created. The market factor and the SMB are the most significant in model performance, while the HML is both redundant and harmful. The factors added by Fama and French (2015) help constituting the model that best explains returns on the 62 portfolios analyzed in this study.

The main limitation of this research lies on the period analyzed. While pricing studies tend to cover the longest period possible, it was not possible to analyze a period longer than 2 years. This limitation is partially compensated by the use of actual data to calculate the VPIN, a rare procedure in market microstructure studies, due to the fact that most scholars do not obtain such data.

As a guide for further research in this area, we suggest deeper analyses of the relation between the factor SMB and the IMU, in order to grasp the depth of their relation. Another gap refers to the extraction of information content from transactions, a trend which is led by Easley, López de Prado e O'Hara (2016) and it still needs a lot of advance, because this is a variable of extreme complexity and volatility.

# REFERENCES

Abad, D., & Rubia, A. (2005). Modelos de estimación de la probabilidad de negociación informada: uma comparación metodológica en el mercado español. *Revista de Economia Financeira*, *7*, 1-37.

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

Agudelo, D., Giraldo, S., & Villaraga, E (2015). Does PIN measure information? Informed trading effects on returns and liquidity in six emerging markets. *International Review of Economics and Finance*, *39*, 149-161.

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

Albanez, T., Lima, G., Lopes, A. & Valle, M. (2010). The relationship of asymmetric information in Brazilian public companies. *Review of Business*, *31*, 3-21.

Albanez, T., & Valle, M. (2009). Impactos da assimetria de informação na estrutura de capital de empresas brasileiras abertas. *Revista de Contabilidade e Finanças*, *20*(51), 6-27.

Barbedo, C., Silva, E., & Leal, R. (2009). Probabilidade de informação privilegiada no mercado de ações, liquidez intra-diária e níveis de governança corporativa. *Revista Brasileira de Economia*, *63*(1), 49-60.

Borochin, P., & Rush, S. (2016). Identifying and pricing adverse selection risk with VPIN (SSRN Working Paper). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2599871

Breeden, D. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economic*, *7*, 265-296.

Brennan, M., Huh, S., & Subrahmanyam, A. (2015). Asymmetric effects of informed trading on the cost of equity capital. *Management Science*, *62*(9), 2460-2480.

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

Chiah, M., Chai, D., Zhong, A., & Li, S. (2016). A better model? An empirical investigation of the Fama-French five-factor model in Australia. *International Review of Finance*, *16*(4), 595-638.

Duarte, J., Hu, E., & Young. L. (2015). What does the PIN model identify as private information? (SSRN Working Paper). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2564369

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

Easley, D., Engle, R., O'Hara, M., & Wu, L. (2008). Time-varying arrival rates of informed and uninformed trades. *Journal of Financial Econometrics*, *6*(2), 171-207.

Easley, D., Hvidkjaer, S., & O'Hara, M. (2002). Is information risk a determinant of asset returns? *The Journal of Finance*, *52*(5), 2185-2221.

Easley, D., Hvidkjaer, S., & O'Hara, M. (2005). Factoring information into returns. *Journal of Financial and Quantitative Analysis*, *45*(2), 293-309.

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. (2011). The microstructure of the "Flash Crash": flow toxicity, liquidity crashes, and the probability of informed trading. *The Journal of Portfolio Management*, *37*(2), 118-128.

Easley, D., López de Prado, M., & O'Hara, M. (2012). Flow toxicity and liquidity in a high frequency world. *Review of Financial Studies*, *25*, 1457-1493.

Easley, D., López de Prado, M., & O'Hara, M. (2016). Discerning information from trade data. *Journal of Financial Economics*, *120*(2), 269-286.

Easley, D., & O'Hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics*, *19*, 69-90.

Fama, E. (1970). Efficient capital markets: a review of theory and empirical work. *The Journal of Finance, 25*(2), 383-417.

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*, 3-56.

Fama, E., & French, K. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, *51*(1), 55-84.

Fama, E., & French, K. (2015). A five-factor asset pricing model. *Journal of Financial Economic*, *116*, 1-22.

Fama, E., & French, K. (2016a). Dissecting anomalies with a five-factor model (SSRN Working Paper). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2503174

Fama, E., & French, K. (2016b). International tests of a five-factor asset pricing model (SSRN Working Paper). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2622782

Fama, E., & French, K. (2017). Choosing factors (SSRN Working Paper). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2668236

Flister, F., Bressan, A., & Amaral, H. (2011). CAPM condicional no mercado brasileiro: um estudo dos efeitos momento, tamanho e *book-to-market*. *Revista Brasileira de Finanças*, *9*(1), 105-129.

Fortunato, G., Motta, L., & Russo, G. (2010). Custo de capital próprio em mercados emergentes: uma abordagem empírica no Brasil com *downside risk*. *Revista de Administração Mackenzie*, *11*(1), 92-116.

Gibbons, M., Ross, S., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, *57*, 1121-1152.

Glosten, L., & Milgrom, P. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, *14*, 71-100.

Gujarati, D. (2006). *Econometria básica*. São Paulo: Campus.

Hwang, L.-S., Lee, W-J., Lim, S-Y., & Park, K-H. (2013). Does information risk affect the implied cost of equity capital? An analysis of PIN and adjusted PIN. *Journal of Accounting and Economics, 55*, 148-167.

Hwang, S., & Satchell, S. (2014). Testing linear factor models on individual stocks using the average F-test. *The European Journal of Finance*, *20*(5), 463-498.

**R. Cont. Fin.** – USP, São Paulo, v. 28, n. 75, p. 425-444, set./dez. 2017

443

Iquiapaza, R., Lamounier, W., & Amaral, H. (2008). Assimetria de informações e pagamento de dividendos na BOVESPA. *Advances in Scientific and Applied Accouting*, *1*(1), 1-15.

Jagannathan, R., & Wang, Z. (1996). The condicional CAPM and the cross-section of expected returns. *Journal of Finance*, *51*(1), 3-53.

Kubota, K., & Takehara, H. (2017). Does the Fama and French five-factor model work well in Japan? *International Review of Finance*, *17*(1), 1-10.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica*, *53*(6), 1315-1336.

Lai, S., Ng, L., & Zhang, B. (2014). Does PIN affect equity prices around the world? *Journal of Financial Economics*, *114*(1), 178-195.

Lambert, R., Leuz, C., & Verrecchia, R. (2011). Informational asymmetry, information precision, and the cost of capital. *Review of Finance*, *16*, 1-29.

Leland, H. (1992). Insider trading: should it be prohibited? *The Journal of Political Economy*, *100*(4), 859-887.

Lewellen, J., & Nagel, S. (2006). The conditional CAPM does not explain asset-pricing anomalies. *Journal of Financial Economics*, *82*, 289-314.

Lin, Q. (2017). Noisy prices and the Fama-French five-factor asset pricing model in China. *Emerging Markets Review*, *30*, 1-23.

Machado, O., Bortoluzzo, A., Sanvicente, A., & Martins, S. (2013). CAPM intertemporal: um teste empírico utilizando dados brasileiros. *Revista Brasileira de Finanças*, *11*(2), 1-18.

Martins, O., & Paulo, E. (2013). A probabilidade de negociação com informação privilegiada no mercado acionário brasileiro. *Revista Brasileira de Finanças*, *11*(2), 249-280.

Martins, O., & Paulo, E. (2014). Assimetria de informação na negociação de ações, características econômico-financeiras e governança corporativa no mercado acionário brasileiro. *Revista Contabilidade e Finanças*, *25*(64), 33-45.

Merton, R. (1973). An intertemporal capital asset pricing model. *Econometrica*, *41*(5), 867-887.

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

Nichol, E., & Dowling, M. (2014). Profitability and investment factors for UK asset pricing models. *Economics Letters*, *125*, 364-366.

O'Hara, M. (1995). *Market microstructure theory*. Cambridge, MA: Blackwell.

Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*, 341-360.

Sanvicente, A., & Bellato, L. (2004). Determinação do grau necessário de diversificação de uma carteira de ações no mercado de capitais brasileiro. In *Seminários em Administração – Semead FEA-USP* (p. 7). São Paulo, SP.

Sharpe, W. (1964). Capital asset prices: a theory of market equilibrium under conditions at risk. *The Journal of Finance*, *19*, 425-442.

Tambosi Filho, E., Garcia, F., Imoniana, J., & Moreiras, L. (2010). Teste do CAPM condicional dos retornos de carteiras dos mercados brasileiro, argentino e chileno, comparando-os com o mercado norte-americano. *Revista de Administração de Empresas*, *50*(1), 60-74.

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

Yildiz, S., Van Ness, R., & Van Ness, B. (2016). The role of HFTs in order flow toxicity and stock price variance, and predicting changes in HFTs liquidity provisions (SSRN Working Paper). Retrieved from https://papers.ssrn.com/sol3/papers2.cfm?abstract_id=2737457

Zaremba, A., & Czapkiewicz, A. (2016). Digesting anomalies in emerging European markets: a comparison of factor pricing models. *Emerging Markets Review*, *29*, 1-15.

## Correspondence address:

**Leonardo Souza Siqueira**
Universidade Federal de Minas Gerais, Faculdade de Ciências Econômicas,
Centro de Pós-Graduação e Pesquisas em Administração
Avenida Presidente Antônio Carlos, 6627 – Sala 4012 – CEP: 31270-901
Pampulha – Belo Horizonte – MG – Brasil
Email: leonardosiq91@gmail.com