

O efeito do risco de informação assimétrica sobre o retorno de ações negociadas na BM&FBOVESPA

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, Brasil

Hudson Fernandes Amaral

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

Universidade de Lisboa, Instituto Superior de Economia e Gestão, Departamento de Gestão, Lisboa, Portugal

Láise Ferraz Correia

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

Recebido em 10.02.2017 – Desk aceite em 01.03.2017 – 2ª versão aprovada em 28.05.2017

RESUMO

Este estudo buscou analisar a assimetria informacional no mercado de ações brasileiro e sua relação com os retornos requeridos de portfólios por meio da métrica *volume-synchronized probability of informed trading*. Para isso, o estudo utilizou dados reais oriundos das transações de 142 ações realizadas na Bolsa de Valores, Mercadorias e Futuros de São Paulo (BM&FBOVESPA), no período de 1º de maio de 2014 a 31 de maio de 2016. Os resultados apontam um alto nível de toxicidade dos fluxos nas ordens dessas ações. Em análises por segmento de listagem na bolsa, os dados sugerem não haver indícios de que as ações dos segmentos teoricamente mais transparentes apresentam menor nível de toxicidade dos fluxos de ordens. A justificativa para esse achado reside na correlação negativa observada entre o valor de mercado das ações e o nível de toxicidade das ordens. Para testar o efeito do risco de informação assimétrica sobre o retorno das ações, um fator relacionado ao nível de toxicidade das ordens foi adicionado aos modelos de três, quatro e cinco fatores. Por meio do teste GRS, verificou-se que a combinação de fatores que otimiza a explicação dos retornos dos portfólios criados foi aquela que se valeu dos fatores mercado, tamanho, lucratividade, investimento e risco informacional. Para testar a robustez desses resultados, empregou-se o Average F-test em dados simulados pelo método *bootstrap*, sendo obtidas estimativas semelhantes. Observou-se que o fator relacionado ao índice *book-to-market* se torna redundante no cenário nacional para os modelos testados. Além disso, verificou-se que o fator relacionado ao risco informacional funciona como um complemento ao fator tamanho e que sua inclusão leva a uma melhora no desempenho dos modelos, indicando um possível poder de explicação do risco informacional sobre o retorno dos portfólios. Os dados sugerem, portanto, que o risco informacional é precificado no mercado acionário brasileiro.

Palavras-chave: risco informacional, retorno de ações, PIN, VPIN, precificação de ativos.

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.

1. INTRODUÇÃO

Devido ao crescente número de ações transacionadas em alta frequência e a concomitante expansão dos bancos de dados *tick by tick*, as pesquisas em microestrutura de mercado tornam-se cada vez mais viáveis. Em especial, isso possibilita que o campo de microestrutura deixe de ser visto somente como um meio de estudar um comportamento de curto prazo dos preços dos ativos e possa, assim, ser associado às demais áreas de estudos em finanças, como a precificação de ativos.

A área de microestrutura de mercado estuda o processo e as consequências da compra e venda de ações (O'Hara, 1995). A principal diferença dessa área para o enfoque tradicional dos modelos de precificação decorre do foco da microestrutura ao analisar como mecanismos específicos de transação afetam a formação do preço das ações. Portanto, uma das vertentes da microestrutura é o estudo do conteúdo informacional carregado pelos preços das ações. A diferença entre as informações que os agentes detêm em um mercado é denominada assimetria informacional e é alvo de estudos pelo menos desde a década de 1970.

Fama (1970) foi um dos pioneiros do estudo do papel do conjunto informacional detido pelos acionistas ao diferenciar a eficiência do mercado em três formas: fraca, semiforte e forte, a depender do modo como o preço do ativo reflete informações sobre ele. Surgiram nesse contexto os modelos de Kyle (1985) e de Glosten e Milgrom (1985), que propõem um dos primeiros modelos de microestrutura de mercado ao considerar os efeitos de *inside trading* sobre os preços de *bid* e *ask*, na perspectiva do formador de mercado.

A partir de Easley e O'Hara (1987), Easley, Kiefer, O'Hara e Paperman (1996) procuraram quantificar a assimetria informacional observada nos preços das ações. A partir de então, diversos estudos, conduzidos principalmente pelos autores supracitados (Easley, Engle, O'Hara, & Wu, 2008; Easley, López de Prado, & O'Hara, 2011), procuraram refinar e desenvolver uma forma de determinação da assimetria informacional observada em um mercado acionário, resultando primeiramente na *probability of informed trading* (PIN) e posteriormente na *volume-synchronized probability of informed trading* (VPIN). A VPIN procura medir diretamente o nível de toxicidade dos fluxos de ordens de uma ação. O termo toxicidade se refere à perda esperada de um *market maker* ao estar no mesmo ambiente de um agente mais bem informado, ou seja, quanto mais tóxico for um fluxo de ordens, maior será a probabilidade de haver um indivíduo

com informações privilegiadas lançando ordens de compra ou venda ao mesmo tempo que outros investidores provêm liquidez, o que resulta em um desequilíbrio de ordens.

Considerando o mercado acionário brasileiro mais arriscado do que os mercados de países desenvolvidos (Martins & Paulo, 2013) e levando em conta que os países emergentes são campos férteis para transações acionadas por agentes com informações privilegiadas (Duarte & Young, 2009), diversos pesquisadores se propuseram a estudar a assimetria informacional no mercado nacional, tanto por meio da PIN (Barbedo, Silva, & Leal, 2009; Martins & Paulo, 2013, 2014) como por modelos alternativos (Iquiapaza, Lamounier, & Amaral, 2008; Albanez & Valle, 2009; Albanez, Lima, Lopes, & Valle, 2010). As evidências empíricas de tais pesquisas convergem para o mesmo resultado: há alta probabilidade da prática de *inside trading* estar presente no mercado acionário brasileiro.

O desequilíbrio informacional acerca de ativos negociados em um mercado financeiro representa risco para os investidores, que demandariam, por consequência, um prêmio para negociar aqueles ativos que percebem como mais arriscados em termos de nível de informação. Assim, o risco informacional de um ativo pode ser um dos fatores precificados pelos agentes em um mercado. O cálculo da taxa de retorno ajustada ao risco das ações gera controvérsia na literatura, tendo surgido diversos modelos que se propõem a mensurá-lo, perpassando os trabalhos de Sharpe (1964), Merton (1973), Jagannathan e Wang (1996), Ross (1976), Fama e French (1993, 2015), dentre outros. Uma dificuldade em sua mensuração reside na determinação dos fatores explicativos que compõem o modelo, sendo o fator de mercado, derivado do modelo de precificação de ativos de capital (CAPM), o mais utilizado para a precificação de ativos no mercado financeiro (Fortunato, Motta, & Russo, 2010).

Apesar de sua extensa utilização na prática e nas pesquisas em finanças, Easley, Hvidkjaer e O'Hara (2005) apontam a incongruência na utilização de modelos como o CAPM para estudar a precificação da assimetria de informações por parte dos investidores. Isso se deve ao fato de que os modelos PIN e VPIN derivam de um cenário em que os participantes possuem diferentes níveis de acesso à informação, portanto, tal fator violaria o pressuposto supracitado, indicando a necessidade da utilização de outros modelos.

Dessa forma, esta pesquisa seguiu os passos de Easley et al. (2005) e Mohanram e Rajgopal (2009), que

estudam a influência do risco informacional sobre o retorno requerido das ações por meio da introdução de um fator relacionado à PIN no modelo de Fama e French (1993). Este estudo vai um passo além ao estimar, por meio dos dados reais, o efeito da VPIN sobre o retorno requerido de ações negociadas no mercado acionário brasileiro, mediante o modelo de 5 fatores de Fama e French (2015). Tendo em vista a necessidade de maiores estudos acerca da existência de investidores informados e sua influência sobre o mercado acionário brasileiro, o objetivo desta pesquisa foi verificar se o nível do fluxo

de ordens das ações, quantificado pela VPIN, é fator de risco sistemático precificado pelos investidores em ações negociadas na Bolsa de Valores, Mercadorias e Futuros de São Paulo (BM&FBOVESPA) entre 1º de maio de 2014 e 31 de maio de 2016.

Este artigo está dividido em 5 seções. A segunda apresenta uma revisão geral da literatura que fundamentou este estudo empírico; a terceira descreve as metodologias utilizadas; e a quarta analisa os resultados dos modelos estimados. A última seção retoma os objetivos e relata nossas considerações finais.

2. REFERENCIAL TEÓRICO

2.1 Assimetria Informacional

O conceito de assimetria informacional é definido por Lambert, Leuz e Verrecchia (2011) como resultante do fato de um conjunto de investidores não ter acesso a informações que estão disponíveis a outros participantes. A utilização de tais informações para compra e venda de ações no mercado financeiro é denominada *inside trading*.

Segundo Leland (1992), inúmeros mercados são caracterizados por ter um desequilíbrio de informações entre compradores e vendedores. Esse fenômeno é ainda mais acentuado nos mercados financeiros, em especial nas relações entre tomadores de recursos e credores. Levando em consideração a hipótese de eficiência de mercado (HEM), de Fama (1970), Leland (1992) aponta os argumentos a favor e contra a prática do *inside trading*. Por um lado, na medida em que os preços das ações refletem todas as informações disponíveis (públicas e privadas), a atuação do *insider* fará com que as novas informações sejam incorporadas aos preços dos ativos. Por outro lado, potenciais investidores serão avessos a entrar nesse mercado por considerá-lo injusto. Assim, os investimentos, preços e liquidez dos ativos serão menores, prejudicando aqueles investidores que operam no mercado sem informações privilegiadas.

Tendo em vista as divergentes conclusões dos pesquisadores acerca do efeito da assimetria informacional, torna-se interessante analisar estatisticamente a pertinência de seu efeito, o que requer um meio de quantificar esse fenômeno.

2.2 Probability of Informed Trading

A PIN foi introduzida por Easley et al. (1996). Primeiro, pressupõe-se que as operações de compra e venda de ativos ocorram com base nas informações que os investidores

possuem. Os autores colocam que o modelo se baseia no fato de que, ao longo do dia, um evento informativo se dá de forma aleatória e é independentemente distribuído, ocorrendo com uma probabilidade α ; δ representa a probabilidade de ser uma notícia ruim e $(1 - \delta)$ uma notícia boa. Negociadores informados operam com uma taxa de negociação μ . Por sua vez, negociadores desinformados transacionam no mercado com taxa de chegada ε_b para compras e ε_s para vendas.

Easley et al. (2005) argumentam que os volumes físicos totais de negociações de compra e venda são suficientes para a estimação da PIN. Dessa forma, os parâmetros do modelo, isto é, o vetor $\theta = (\alpha, \mu, \varepsilon_b, \varepsilon_s, \delta)$, podem ser estimados a partir da maximização de uma função de máxima verossimilhança.

Como se trata de um modelo probabilístico que envolve a ocorrência de um conjunto de diferentes eventos relacionados, a probabilidade de negociação com informação privada (PIN) segue a formulação (1):

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

Por ser um modelo que utiliza dados intradiários e avalia diretamente a probabilidade da atuação de agentes informados, a PIN tem sido largamente utilizada na literatura financeira e foi testada empiricamente nos mercados norte-americano (Easley et al., 2005; Easley, Hvidkjaer, & O'Hara, 2002); espanhol (Abad & Rubia, 2005); brasileiro (Barbedo et al., 2009; Martins & Paulo, 2013, 2014; Agudelo, Giraldo, & Villaraga, 2015); francês (Aktas, Bodt, Declerck, & Van Oppens, 2007); sul-coreano (Hwang, Lee, Lim, & Park, 2013); colombiano, argentino, chileno, peruano e mexicano (Agudelo et al., 2015), dentre outros.

Duarte e Young (2009) apontam que a PIN condensa os motivos que levam os investidores a lançar ordens de transações em somente dois: informações privilegiadas ou busca por liquidez. Em estudo mais recente, Duarte, Hu e Young (2015) sugerem que a PIN não consegue capturar informações privilegiadas devido à sua mecânica de operação.

No mercado brasileiro, Barbedo et al. (2009) estudaram a relação da PIN com os níveis de governança corporativa da BM&FBOVESPA. Já Martins e Paulo (2013, 2014) aplicaram a PIN no mercado brasileiro nos períodos de 2010 e 2011, procurando relacionar o resultado encontrado com os níveis de governança corporativa e as características econômico-financeiras das empresas, como: risco, retorno, liquidez, custo de capital e tamanho das firmas, dentre outras. Os autores encontram uma probabilidade média de transações privilegiadas de 25% para as empresas no período analisado, valor maior do que o encontrado por Barbedo et al. (2009), de 12,5%.

Devido às diversas críticas à PIN, Easley et al. (2011) desenvolveram um novo modelo de estimação da toxicidade dos fluxos das ordens de transações, denominada VPIN.

2.3 Volume-Synchronized Probability of Informed Trading

Apesar da extensa aplicação empírica da PIN e de sua relevância no campo das finanças – ver revisão de Mohanram e Rajgopal (2009) –, seu cálculo apresenta problemas como a não convergência da função de máxima verossimilhança para dias em que o número de ordens é alto. Dessa forma, Easley, López de Prado e O'Hara (2012) propuseram o *volume-synchronized probability of informed trading* (VPIN) com base em Easley et al. (2008). Essa métrica, além de sanar o problema mencionado, busca quantificar diretamente o nível de toxicidade dos fluxos de ordens sem a necessidade da estimação de parâmetros por meio de funções de máxima verossimilhança.

Easley et al. (2008) mostram que o valor esperado para a soma dos volumes de compra e venda é igual à quantidade total transacionada, representada pelo denominador da equação (1). Ao mesmo tempo, a diferença entre o volume de compra e venda pode ser aproximada ao valor da taxa de negociadores informados vezes a probabilidade de ter ocorrido um evento informacional, ou seja, o numerador de (1). Essas relações são representadas pelas equações (2) e (3):

$$E[V_t^B + V_t^S] = \alpha\mu + 2\varepsilon = nV \quad 2$$

$$E[V_t^B - V_t^S] \approx \alpha\mu \quad 3$$

A ideia da VPIN reside na divisão do dia em baldes de volumes iguais (*volume buckets*), tratando cada um equivalente a um período de chegada de informação. Dessa forma, $V_\tau^B + V_\tau^S$ é constante e igual a V para todo τ .

Aproxima-se, então, o desequilíbrio de transações pela média calculada sobre baldes de volume. Dessa forma, a VPIN pode ser calculada por (4):

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

2.4 Modelos de Precificação de Ativos

Diversos modelos de precificação de ativos foram propostos na literatura ao longo dos anos, sendo um dos primeiros e mais impactantes o CAPM, de Sharpe (1964). O CAPM pressupõe que o retorno esperado dos ativos é uma função linear do seu beta, multiplicado pelo prêmio de risco de mercado somado a um retorno de um ativo livre de risco.

Derivados do CAPM, outros trabalhos foram desenvolvidos visando a aumentar a robustez do modelo original, como o CAPM intertemporal (ICAPM), de Merton (1973), o CAPM baseado em consumo, de Breeden (1979), e o CAPM condicional (C-CAPM), de Jagannathan e Wang (1996). Testes em mercados internacionais mostram que o C-CAPM não foi capaz de explicar anomalias nos retornos dos ativos (Lewellen & Nagel, 2006). No mercado brasileiro, Tambosi Filho,

Garcia, Imoniana e Moreiras (2010) e Flister, Bressan e Amaral (2011) não encontram evidências contrárias ao C-CAPM, porém, recomendam cautela em sua utilização, principalmente devido à imaturidade do mercado nacional. Machado, Bortoluzzo, Sanvicente e Martins (2013) analisam a aplicação do ICAPM no Brasil e verificam que os resultados são favoráveis ao modelo para o período entre 2003 e 2011.

As evidências se mostraram, portanto, inconclusivas quanto ao modelo mais adequado para precificar ações. Com a evolução dos estudos de precificação de ativos, outras fontes de risco foram incorporadas à explicação dos retornos das ações. Fama e French (1993), por exemplo, encontram que existem pelo menos três fatores que afetam os retornos dos ativos analisados. Estes são os conhecidos *small minus big* (SMB), *high minus low* (HML) e o fator de mercado, o que resultou no modelo de 3 fatores de Fama e French (1993).

$$R_i - R_f = \alpha_i + b_iMKT + s_iSMB + h_iHML + r_iRMW + c_iCMA + \varepsilon_i \quad \boxed{5}$$

Fama e French (2015) fornecem várias contribuições em relação ao seu modelo anterior, por exemplo, a possibilidade da criação dos fatores por meio de combinações diferentes daquela usada em Fama e French (1993, 1996). De forma a corroborar a decisão pela utilização do modelo (5), Fama e French (2015) mostram que o valor da estatística de Gibbons, Ross e Shanken (1989) – GRS – é menor para o modelo de 5 fatores do que para o de 3 fatores.

Fama e French (2016a) colocam que as anomalias referentes à aplicação do CAPM diminuem quando se aplica o modelo de 5 fatores. Além disso, este último conseguiu sanar problemas referentes ao modelo de 3 fatores, como aqueles relacionados à recompra de ações. Em testes em mercados internacionais, Fama e French (2016b) atestam a superioridade do modelo de 5 fatores em relação aos demais, atentando, porém, às falhas relacionadas à explicação dos retornos de ações de pequena capitalização. Outros trabalhos em mercados como o australiano (Chiah, Chai, Zhong, & Li, 2016), japonês (Kubota & Takehara, 2017), chinês (Lin, 2017), inglês (Nichol & Dowling, 2014), além de um trabalho que agrega diferentes mercados de países europeus (Zaremba & Czapkiewicz, 2016), mostram a superioridade desse modelo em relação aos demais.

Fama e French (2017) testam variações dos fatores propostos por Fama e French (2015) e mostram que a escolha dos fatores são respostas aos problemas empíricos

Apesar do aparente sucesso do modelo de 3 fatores em comparação ao CAPM, Fama e French (1996) atentam que ele não é capaz de explicar os retornos de todos os ativos e portfólios. Dessa forma, Carhart (1997) sugere a adição de um quarto fator, denominado fator de momento *up minus down* (UMD).

Devido a evidências surgidas na literatura ao longo dos anos de que os modelos de 3 e 4 fatores não são capazes de explicar a variação nos retornos médios relativos à lucratividade e ao investimento, Fama e French (2015) revisitam seu modelo anterior, adicionando a ele 2 fatores. O primeiro deles, denominado *robust minus weak* (RMW), é obtido pela diferença entre os retornos de portfólios de ações com alta e baixa lucratividade. Já o fator referente ao investimento, denominado *conservative minus aggressive* (CMA) é a diferença entre o retorno de portfólios de ações com baixo e alto investimento. O modelo em sua forma completa é representado por (5):

do CAPM e do C-CAPM. A escolha dos fatores se relaciona, portanto, à descoberta de padrões nos retornos dos ativos e, na medida em que tais padrões mudam ao longo do tempo, novos fatores podem ser adicionados aos modelos.

Baseando-se nos modelos de precificação com solidez na literatura, diversos autores buscaram uma relação entre a probabilidade de haver negociações privilegiadas em um mercado e o retorno requerido das ações. Os resultados encontrados pelos autores apresentam divergências, tendo, em alguns trabalhos, sido constatada relação positiva entre a PIN e o retorno requerido e, em outros, inexistência de relação entre eles.

2.5 PIN e os Modelos de Precificação de Ativos

Dentre os trabalhos que procuraram incorporar um fator de risco informacional aos modelos de precificação de ativos, destacam-se Easley et al. (2002, 2005), Mohanram e Rajgopal (2009) e Hwang et al. (2013), onde os pesquisadores se propuseram a analisar empiricamente a influência da PIN sobre o retorno requerido de ações negociadas no mercado acionário norte-americano. Enquanto Easley et al. (2002, 2005) e Hwang et al. (2013) afirmam que o risco informacional é precificado pelos investidores por meio de um fator de risco sistemático a ele relacionado, Mohanram e Rajgopal (2009) encontraram resultados contraditórios a essa afirmação.

Easley et al. (2005) criam um fator PIN e o adicionam ao modelo de 3 fatores de Fama e French (1993) e ao modelo de 4 fatores de Carhart (1997). Os resultados encontrados apontam um intercepto estatisticamente insignificante para 8 dos 10 portfólios quando acrescido o fator PIN na regressão. Em posse dos dados reais, Hwang et al. (2013) calculam a PIN e a regridem com o retorno esperado representado por 4 estimativas diferentes do custo de capital próprio implícito. Os autores chegam a resultados empíricos que sustentam a hipótese de que existe uma relação entre risco informacional e os retornos esperados, assim como relatado por Easley et al. (2005).

Mohanram e Rajgopal (2009) replicando os trabalhos de Easley et al. (2002) concluem que os retornos para o fator PIN são negativamente correlacionados com os retornos de ações com alta PIN. Além disso, o fator PIN não apresentou um coeficiente significativo no teste com os modelos de 3 e 4 fatores de Fama e French (1993) e Carhart (1997).

Além desses trabalhos, Brennan, Huh e Subrahmanyam

(2015) encontram evidências da precificação da assimetria informacional no mercado norte-americano por meio da decomposição da PIN em 2 fatores. Borochin e Rush (2016), utilizando a VPIN para a criação de um fator de precificação, encontram resultados favoráveis à hipótese de que há um efeito relacionado ao risco informacional precificado pelos agentes. Lai, Ng e Zhang (2014) apresentam evidências contrárias ao poder de explicação do fator PIN ao analisar ações de 47 países. Assim como Duarte e Young (2009), esses autores concluem que a PIN pode estar mais relacionada à mudança na demanda por liquidez das ações do que seu conteúdo informacional.

Os resultados desses estudos apontam uma necessidade ainda mais profunda de analisar a relação do risco informacional e o retorno requerido de ações. No Brasil, por exemplo, Martins e Paulo (2014) encontraram uma relação positiva da PIN com o custo de capital e retorno das ações. Esses autores, porém, não se valeram dos dados reais das ordens de transação, havendo, portanto, a possibilidade de problemas com a classificação das ordens.

3. METODOLOGIA

3.1 População e Amostra

O objeto deste estudo consistiu em todas as ações transacionadas na BM&FBOVESPA. Para o cálculo da VPIN, a amostra se restringiu àquelas ações que apresentaram no mínimo uma transação por dia entre 1º de maio de 2014 e 31 de maio de 2016, período no qual os dados estavam disponíveis no *market data* da BM&FBOVESPA. Dessa forma, o número de ativos disponíveis para o cálculo dessa variável foi de 142 ações (ordinárias e preferenciais). Para as formações dos fatores de Fama e French (1993, 2015) e Carhart (1997), a amostra disponível foi de 349 ações.

Na amostra, analisaram-se tanto ações preferenciais como ordinárias de uma mesma empresa, uma vez que a evidência empírica no Brasil indica que elas carregam conteúdos informacionais diferentes. Martins e Paulo (2013) encontraram menores valores médios de PIN para ações ordinárias em relação aos das preferenciais, mesmo após considerar o nível de liquidez das ações. Neste estudo, no entanto, não foi atribuída ênfase à diferença da VPIN para as diferentes classes de ações.

3.2 Coleta e Tratamento dos Dados

A principal limitação dos trabalhos que se propuseram a aplicar a PIN ou, neste caso, a VPIN, relaciona-se à

classificação errônea das ordens de compra e venda. De forma a contornar o problema, esta pesquisa se vale dos dados reais do volume transacionado no mercado brasileiro. Não há registro, na literatura brasileira, de pesquisa que utilize o *market data* da BM&FBOVESPA para o cálculo da VPIN, o que evidencia a originalidade deste trabalho na área de microestrutura de mercado em finanças.

O principal motivo da utilização pouco frequente desse diretório pode residir na dificuldade do tratamento dos dados. Estes são disponibilizados via arquivos de texto, contendo inúmeras informações, como: preço do negócio, quantidade transacionada, horário, condição da oferta, código das corretoras envolvidas, indicador do tipo de ordem, compra ou venda, dentre outros dados. Assim, lidar com esses arquivos requer um árduo e longo trabalho, além da necessidade de maior poder computacional para a separação e filtragem das informações relevantes para a aplicação dos modelos.

Os dados coletados por meio do *market data* da BM&FBOVESPA foram tratados exclusivamente em rotinas desenvolvidas no programa estatístico R. As diversas outras informações, como excesso de retorno da ação, tamanho, *book-to-market*, lucratividade, ativo, passivo não circulante e demais dados necessários para a aplicação dos modelos foram coletados das bases de dados Bloomberg® e Quantum Axis®.

3.3 Hipóteses de Pesquisa

Para atingir os objetivos desta pesquisa, algumas hipóteses foram testadas em relação às variáveis estudadas; são elas:

Hipótese 1: Quanto menor o tamanho da empresa, maior o nível de toxicidade dos fluxos das ordens.

Segundo Easley et al. (1996), ações de grandes empresas têm maior cobertura de analistas e, também, maior atenção dos investidores. Dessa forma, a chance de haver transações privilegiadas é, em tese, menor para essas ações, resultando em um VPIN mais baixo do que o encontrado para ações de companhias menores, assim como verificado também por Abad e Yagüe (2012) e Wei, Gerace e Frino (2013).

Hipótese 2: Os segmentos de listagem da BM&FBOVESPA apresentam VPIN diferenciados.

Esperava-se que as empresas que compõem

os segmentos diferenciados da BM&FBOVESPA apresentassem menor VPIN, assim como verificado por Barbedo et al. (2009) e Martins e Paulo (2013). Dessa forma, estipulou-se a hipótese de que a VPIN para o segmento NM é o mais baixo, seguido do N2, N1 e, por fim, o tradicional.

Hipótese 3: Um fator relacionado à VPIN ajuda a explicar os retornos de portfólios.

Esperava-se que a adição de um fator VPIN nos modelos de 3 e 5 fatores de Fama e French (1993, 2015) e de quatro fatores de Carhart (1997) apresentasse uma redução no intercepto geral das carteiras analisadas pelo testes GRS e Average F-test de Hwang e Satchell (2014).

3.4 Modelos de Fatores e a VPIN

A Tabela 1 apresenta as especificações dos modelos estimados neste estudo, os quais foram baseados em Easley et al. (2005) e Mohanram e Rajgopal (2009).

Tabela 1 Modelos testados e suas respectivas bases empíricas

Modelo	Especificação do modelo	Base empírica
Três Fatores	$R_i - R_f = \alpha_i + \beta_i MKT + s_i SMB + h_i HML$	Fama e French (1993)
Três Fatores + IMU	$R_i - R_f = \alpha_i + \beta_i MKT + s_i SMB + h_i HML + p_i IMU_i$	Fama e French (1993)
Quatro Fatores	$R_i - R_f = \alpha_i + \beta_i MKT + s_i SMB + h_i HML + u_i UMD$	Carhart (1997)
Quatro Fatores + 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)
Cinco Fatores	$R_i - R_f = \alpha_i + \beta_i MKT + s_i SMB + h_i HML + r_i RMW_i + c_i CMA$	Fama e French (2015)
Cinco Fatores + 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 e French (2015)

Fonte: Elaborada pelos autores.

3.5 Variáveis Dependentes

A variável dependente consiste no excesso de retorno diário médio dos portfólios de ações em relação ao CDI [$R_i - R_f$], formados conforme o procedimento adotado por Fama e French (2015). Para a criação dos portfólios manteve-se a variável principal ‘tamanho’ e permutou-

se o segundo componente entre as demais variáveis: *book-to-market*, lucratividade, investimento e VPIN. Seguindo Sanvicente e Bellato (2004), os portfólios com menos de 6 ações foram excluídos dos testes, por não ser suficientemente diversificados. A Tabela 2 resume esse procedimento e ilustra as informações dos portfólios criados.

Tabela 2 Informações acerca dos portfólios criados a partir da interseção das ações divididas em grupos baseados nas variáveis: tamanho, book-to-market, investimento, lucratividade e VPIN

Variáveis usadas na divisão das ações para criação dos portfólios	Quantidade de portfólios	Mínimo de ações	Máximo de ações	Média de ações
Tamanho (3) / Book-To-Market (4)	12	6	25	14
Tamanho (3) / Investimento (4)	12	7	23	14
Tamanho (2) / Investimento (3) / Lucratividade (2)	12	7	18	13
Tamanho (3) / VPIN (4)	10	6	29	11
Tamanho (4) / Lucratividade (4)	16	6	15	10
Total	62	6	29	12

A coluna 'Variáveis' indica quais as variáveis usadas na construção dos portfólios. Os valores entre parênteses se referem aos breakpoints usados na divisão das ações, por exemplo, (3) indica que as ações foram divididas em 3 grupos. Foram analisadas as interseções entre as ações segundo os valores das variáveis e, assim, formaram-se os portfólios. A quantidade de portfólios é, portanto, a multiplicação dos breakpoints para cada variável. Atenta-se ao fato de que 2 carteiras foram excluídas da combinação Tamanho e VPIN.

Fonte: Elaborada pelos autores.

3.6 Variáveis Independentes

Para a criação de um fator relacionado ao risco informacional, primeiro, as ações foram divididas

em 3 grupos com base em seus valores de mercado. Paralelamente, dividiram-se as ações em 2 grupos: baixa e alta VPIN. Finalmente, calculou-se o retorno ponderado de cada interseção (Tabela 3).

Tabela 3 Portfólios formados com base nas variáveis tamanho e VPIN para criação do fator IMU

Carteira	Sigla	Descrição
Small e Low	SL	Interseção entre as ações do grupo <i>small</i> para a variável tamanho e as ações com baixo valor de VPIN
Small e High	SH	Interseção entre as ações do grupo <i>small</i> para a variável tamanho e as ações com alto valor de VPIN
Medium e Low	ML	Interseção entre as ações do grupo <i>medium</i> para a variável tamanho e as ações com baixo valor de VPIN
Medium e High	MH	Interseção entre as ações do grupo <i>medium</i> para a variável tamanho e as ações com alto valor de VPIN
Big e Low	BL	Interseção entre as ações do grupo <i>big</i> para a variável tamanho e as ações com baixo valor de VPIN
Big e High	BH	Interseção entre as ações do grupo <i>big</i> para a variável tamanho e as ações com alto valor de VPIN

Fonte: Elaborada pelos autores.

Assim, o fator IMU foi obtido como mostrado em (6):

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

6

Os motivos que deram suporte à criação do IMU repousam na relação entre a probabilidade de negociações privilegiadas e o retorno das ações. Easley et al. (2002) encontram correlação positiva entre essas 2 variáveis. Segundo esses autores, ações com maior PIN apresentam um retorno requerido maior e, conseqüentemente, maior custo de capital. Neste estudo, observou-se que a correlação entre a VPIN e o retorno foi de 0,0141, com $p = 0,0003$. Apesar do baixo valor, supõe-se possível a existência de um prêmio para o investimento em ações com maior toxicidade dos fluxos das ordens.

Para a criação dos demais fatores, seguiram-se os procedimentos realizados por Fama e French (1993, 2015)

e Carhart (1997). O fator SMB utilizado nos modelos de 3 e 4 fatores de Fama e French (1993) e Carhart (1997) foi calculado tomando o retorno da interseção das ações com baixo valor de mercado e baixo, médio e alto *book-to-market* e subtraindo o retorno da interseção das ações com alto valor de mercado e baixo, médio e alto *book-to-market*. O fator HML foi criado a partir do retorno da interseção das ações com alto *book-to-market* e os grupos baixo e alto valor de mercado subtraído do retorno da interseção das ações com baixo *book-to-market* e baixo e alto valor de mercado. Obteve-se o fator de momento tomando-se o retorno da interseção das ações com alto retorno passado e baixo e alto valor de mercado, subtraído

do retorno da interseção das ações com baixo retorno passado e baixo e alto valor de mercado.

Para o fator de tamanho utilizado nos modelos de 5 fatores de Fama e French (2015), tomou-se o retorno da interseção das ações com baixo valor de mercado e os demais fatores, subtraído do retorno da interseção das ações com alto valor de mercado e os demais fatores. O fator HML foi criado a partir do retorno da interseção das ações com alto *book-to-market* e baixo e alto valor de mercado subtraído do retorno da interseção das ações com baixo *book-to-market* e baixo e alto valor de mercado.

Para o fator relacionado à lucratividade (RMW), tomou-se o retorno da interseção das ações com alta lucratividade e baixo e alto valor de mercado, subtraído do retorno da interseção das ações com baixa lucratividade e baixo e alto valor de mercado. Finalmente, para o fator de investimento (CMA), o retorno da interseção das ações com baixo investimento e baixo e alto valor de mercado foi subtraído do retorno da interseção das ações com alto

investimento e baixo e alto valor de mercado.

3.7 Simulação dos Portfólios por *Bootstrap*

A simulação por *bootstrap* foi utilizada para analisar qual combinação de fatores otimiza a explicação dos retornos dos portfólios. A aplicação desse método consistiu em selecionar os retornos dos portfólios criados por meio de reamostragem com reposição, e sob cada uma delas, estimaram-se os modelos, calculando, portanto, os coeficientes de cada regressão. A partir da obtenção dos interceptos das regressões, aplicou-se o Average F-test. A aplicação desse teste se fez necessária, pois a simulação dos dados gera dependência linear entre eles, tornando impossível a inversão da matriz de covariância dos resíduos das regressões, necessária para a utilização do teste de GRS. Para os momentos em que não se utilizou dados simulados, o teste GRS foi aplicado seguindo os passos de Fama e French (1996, 2015).

4. ANÁLISE DOS RESULTADOS

4.1 Resultados da VPIN

O resultado para o cálculo da VPIN é apresentado na Tabela 4. A comparação dos resultados de Barbedo et al. (2009) e Martins e Paulo (2013, 2014) encontra dificuldades inerentes aos procedimentos utilizados em

cada pesquisa. Barbedo et al. (2009) e Martins e Paulo (2013, 2014) aplicaram a PIN, além de ter utilizado algoritmos de classificação das ordens de compra e venda. Dessa forma, será apresentada uma comparação de estudos em outros mercados que se valeram da VPIN.

Tabela 4 Estatísticas descritivas da VPIN para toda a amostra entre 01/05/2014 e 31/05/2016

Número de ações	Mínimo	Máximo	Média	Desvio padrão
142	0	1	0,4548	0,2218

Fonte: Elaborada pelos autores.

4.2 VPIN por Segmento de Listagem da BM&FBOVESPA

Em relação aos segmentos de listagem diferenciados da BM&FBOVESPA esperava-se que empresas que se

encontram no segmento Novo Mercado (NM), que têm, em tese, maior transparência do que aquelas dos segmentos Nível 1 (N1), Nível 2 (N2) e Tradicional (Trad), apresentassem menor VPIN. Para analisar essa hipótese, a VPIN diária foi calculada para cada segmento (Tabela 5).

Tabela 5 Estatísticas descritivas da VPIN diária por segmento de listagem da BM&FBOVESPA

Estatística descritiva	Segmentos de listagem			
	N1	N2	NM	Trad
Número de ações	24	12	94	12
Mínimo	0,3168	0,3669	0,4270	0,3466
Máximo	0,4310	0,6001	0,5039	0,5695
Média	0,3761	0,4870	0,4700	0,4597
Desvio padrão	0,0191	0,0355	0,0124	0,0379

Fonte: Elaborada pelos autores.

Verifica-se que a VPIN do Nível 2 foi o maior entre os segmentos de listagem da BM&FBOVESPA, seguido pelo Novo Mercado, Tradicional e, finalmente, o Nível 1. Vale notar que a quantidade de ações em cada segmento difere bastante, sendo o Novo Mercado o segmento com maior número de empresas. Tal fato impacta fortemente os resultados por segmento, como será explicado na análise por valor de mercado.

De forma a verificar estatisticamente a diferença entre as médias da VPIN dos segmentos, utilizou-se o teste *t* de Student. Os resultados apontam a rejeição da hipótese nula de igualdade de valores para a média da VPIN para todos os segmentos. Esperava-se que os valores apresentados pelo NM fossem significativamente mais baixos do que os dos outros segmentos, em especial do Tradicional. Os resultados encontrados foram, no entanto, contrários ao esperado. Constatou-se que o segmento N1 apresentou o menor VPIN para a amostra analisada e que o segmento N2 exibiu maior VPIN médio no período em questão.

4.3 Análise da VPIN por Valor de Mercado da Ação

Um dos principais resultados dos estudos que aplicaram a PIN e, mais recentemente, a VPIN, é de que

a probabilidade de negociação privilegiada é menor para ações de empresas maiores. Para a VPIN, por exemplo, a disparidade entre os volumes de compra e de venda não são tão acentuados para empresas de maior capitalização, reduzindo, portanto, seu nível de toxicidade dos fluxos. Abad e Yagüe (2012) foram os primeiros a constatar tal relação em ações transacionadas no mercado espanhol. Wei et al. (2013) e Yildiz, Van Ness e Van Ness (2016) encontraram evidências que dão suporte a tal afirmação em estudos nos mercados australianos e norte-americanos, respectivamente. Com isso em vista, busca-se verificar a relação entre a VPIN e o valor das empresas no mercado acionário brasileiro.

Para investigar essa relação, dividiram-se as ações em 3 grupos denominados *pequeno*, *médio* e *grande*, com cerca de 47 ações cada, relacionados ao seu valor de mercado diário médio. O primeiro indício relacionando a VPIN com o tamanho das empresas veio com o cálculo da correlação entre essas 2 variáveis. O resultado para a correlação foi de -0,3080 e $p = 0$. Tal valor era esperado, tendo em vista as constantes evidências empíricas da relação negativa entre a VPIN e o tamanho. Com vistas a aprofundar a análise dessa relação, procedeu-se ao cálculo das estatísticas descritivas para cada grupo (Tabela 6).

Tabela 6 Estatísticas descritivas da VPIN diário por grupo, segundo a capitalização de mercado das ações

Estatísticas descritivas	Grupos		
	Pequeno	Médio	Grande
Mínimo	0,5582	0,3544	0,2793
Máximo	0,7134	0,4565	0,3528
Média	0,6364	0,3989	0,3164
Desvio padrão	0,0284	0,0164	0,0123

Fonte: Elaborada pelos autores.

A Tabela 6 evidencia a relação verificada nos trabalhos que aplicaram a VPIN na comparação entre as capitalizações de mercado das ações. Observa-se grande diferença entre a VPIN das ações de menor capitalização

da amostra em relação às de média e grande capitalização. Para efeito de comparação, a Tabela 7 ilustra os resultados obtidos em outros mercados onde a VPIN foi aplicada.

Tabela 7 Comparação entre a VPIN calculada por grupos de tamanho em diferentes mercados

Mercados	Grupos			Referências
	Pequeno	Médio	Grande	
Brasil	0,6364	0,3989	0,3164	Este estudo
Espanha	0,5320	0,4020	0,2540	Abad e Yagüe (2012)
Austrália	0,8459	0,7249	0,5002	Wei et al. (2013)
Estados Unidos	0,6154	0,4001	0,2378	Yildiz et al. (2016)

Fonte: Elaborada pelos autores.

De forma geral, os resultados obtidos para as empresas médias e grandes nos mercados espanhol e norte-americano foram próximos aos do mercado brasileiro. Pode-se destacar que a VPIN das ações de empresas brasileiras apresenta comportamento similar à das empresas dos mercados citados, com correlação negativa ao tamanho das companhias. De forma geral, a VPIN calculada para o mercado nacional, de 0,4548, não se diferencia tanto daquela dos mercados espanhol (0,3960) e norte-americano (0,4178).

Voltando à análise das características da VPIN em relação ao tamanho das empresas, a Figura 1 mostra o comportamento da VPIN de cada grupo. Pode-se observar a diferença substancial do nível de toxicidade das ações em relação a seus valores de mercado. O grupo *pequeno* apresentou um VPIN diário sempre acima de 0,55, enquanto o grupo *médio* ficou em torno de 0,40, com leves picos chegando a 0,45. O grupo *grande* se manteve mais estável, com o menor desvio padrão entre os 3, apresentando um máximo de 0,35.



Figura 1 Evolução da VPIN diário para os grupos de tamanho das ações.

P: grupo de ações pequenas; M: ações médias; G: ações grandes em valor de mercado.

Fonte: Elaborada pelos autores.

Constatou-se uma diferença substancial da VPIN das ações em relação à variável tamanho, a explicação para os comportamentos da VPIN para os segmentos da BM&FBOVESPA pode estar contida nos valores de mercado das ações que compõem cada segmento. Das 47 empresas do grupo *pequeno*, 36 se encontram no segmento Novo Mercado, que tem 94 ações no total, fazendo com que a VPIN médio do NM aumente substancialmente. Excluindo as ações do grupo *pequeno*, a VPIN do NM cairia para 0,3418, valor significativamente menor do que a VPIN atual do segmento. Para o segmento N1, dentre as 24 ações que o compõem, 1 está dentro do grupo *pequeno*, 11 dentro do *médio* e 12 dentro do grupo das grandes empresas, fazendo com que a VPIN seja levada para baixo, o que explicaria o fato do N1 apresentar a menor VPIN entre os segmentos. Caso as ações do grupo *pequeno* e *médio* fossem excluídas do cálculo da VPIN do N1, sua

VPIN cairia de 0,3761 para 0,2972.

Em relação ao segmento N2, das 12 ações que o compõem, 6 são do grupo *pequeno*, 3 do grupo *médio* e 3 do grupo *grande*. Como a VPIN não considera a ponderação pelo tamanho da empresa e dá pesos iguais para as empresas, o fato de metade do segmento ser composta de ações do grupo *pequeno* explicaria por que o N2 apresenta o maior VPIN entre os segmentos. Já para o segmento Tradicional, das 12 ações que o compõem, 3 são do grupo *pequeno*, 3 do grupo *médio* e 6 do grupo *grande*. Dentre as ações do grupo de maior valor de mercado se encontram os seguintes ativos: ABEV3, LAME3, LAME4, PETR3, PETR4 e VIVT4. Juntas, essas ações apresentariam um VPIN médio de 0,3030, bastante inferior aos 0,4597 apresentados pelo segmento como um todo. A razão para tal diferença reside na VPIN das ações menores, que apresentam um valor médio de 0,7835, fazendo com que o

valor da VPIN do segmento Tradicional suba para 0,4597.

A hipótese, portanto, de que o tamanho da empresa e sua VPIN estão negativamente correlacionados foi verificada a partir da amostra analisada no mercado acionário nacional. Tal evidência está em consonância com o esperado e com o apresentado na literatura internacional.

4.4 Análise dos Modelos de 3, 4 e 5 Fatores e o Fator Baseado na VPIN (IMU)

O primeiro passo para analisar o desempenho dos modelos se refere à correlação entre seus fatores (Tabela 8). Percebe-se, pelo painel (a) da Tabela 8, que o IMU apresentou correlação moderada e negativa com o fator de mercado, sendo esse resultado próximo ao reportado por Easley et al. (2005) e Mohanram e Rajgopal (2009). Em relação aos fatores HML e UMD, o IMU não apresentou correlação significativa, com $p = 0,1643$ e $0,9638$.

Verifica-se que a correlação de maior peso para o fator construído pela VPIN é com o SMB. É possível que os retornos das ações com maior VPIN sejam daquelas empresas menores, enquanto as ações com menor VPIN representam as ações maiores, sendo a explicação para a forte correlação entre os 2 fatores sua forma de construção.

Quando analisadas as correlações entre os 5 fatores de Fama e French (2015) e o IMU – painel (b) da Tabela 8 –, verifica-se que, mesmo em diferentes metodologias, o IMU apresenta uma correlação ligeiramente maior com o fator tamanho. Outra correlação moderada e positiva surge entre o IMU e o CMA. De forma geral, os resultados não se assemelham àqueles apresentados por Fama e French (2015), exceto para a correlação positiva entre os fatores CMA e HML. Novamente, os fatores MKT e SMB apresentam correlação moderada e negativa, ao passo que para Fama e French (2015) há uma correlação com a mesma magnitude, porém positiva.

Tabela 8 Correlação entre os 3, 4 e 5 fatores de Fama e French (1993, 2015) e Carhart (1997) e o fator IMU

Painel (a): Correlação entre os 3 fatores de Fama e French (1993), o fator de momento de Carhart (1997) e o fator 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	
Painel (b): Correlação entre os 5 fatores de Fama e French (2015) e o fator 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

Fonte: Elaborada pelos autores.

4.4.1 Resultados das regressões dos modelos de fatores.

Para apresentar os resultados referentes às regressões estimadas, mostra-se necessária a aplicação de testes econométricos que visam a testar a robustez estatística dos modelos analisados. Foram realizados 3 testes para verificar, respectivamente, se existe multicolinearidade, se os resíduos das regressões estão autocorrelacionados e se estes são heterocedásticos. O primeiro teste, *variance inflation factor* (VIF), testa se as variáveis explicativas são correlacionadas, o que afetaria a estimação de seus coeficientes. Seguindo Gujarati (2006), valores de VIF acima de 10 indicam multicolinearidade. Verificou-se que, em nenhum dos 6 modelos, o resultado para o VIF das variáveis foi elevado. Os resultados, portanto, indicam que não há multicolinearidade entre as variáveis, sendo

possível sua inclusão em um modelo de regressão sem aparente prejuízo nas estimativas dos coeficientes.

Foram aplicados também os testes de Durbin-Watson e Breusch-Pagan. Em relação ao primeiro teste, dos 62 portfólios, apenas 1 rejeitou a hipótese nula de ausência de autocorrelação dos resíduos, tendo apresentado estatística de teste de 1,77, ou seja, a correlação foi positiva, porém, não se mostrou forte. Em relação ao teste de Breusch-Pagan, apenas 6 portfólios rejeitaram a hipótese nula de homocedasticidade. Tendo em vista esses resultados, conclui-se que os estimadores MQO são suficientes para a estimação concisa dos coeficientes dos modelos.

A Tabela 9 mostra os resultados para o teste GRS aplicado nos conjuntos de portfólios. Verifica-se que, para os portfólios construídos com base no tamanho e

book-to-market, a adição do fator IMU aos modelos de 3, 4 e 5 fatores leva a melhora em seus desempenhos. Para as carteiras formadas pelo tamanho e investimento das empresas, a adição do IMU não acarreta diferenças significativas em relação aos modelos tradicionais. Um detalhe a destacar é que o modelo de 5 fatores apresentou desempenho pior do que os de 3 e de 4 fatores para esse conjunto de portfólios.

O que se pode verificar é que não existe um modelo que obteve melhor desempenho independentemente do conjunto analisado. Fama e French (2015) mostram que o modelo de 5 fatores apresenta melhora substancial diante do modelo de 3 fatores para os 7 conjuntos de carteiras por eles analisados. Neste estudo, o modelo de

5 fatores foi superior naqueles conjuntos criados a partir das variáveis tamanho e VPIN e tamanho e lucratividade. A formação dos portfólios com base na lucratividade apresentou o melhor grau de explicação por parte dos modelos em geral. Quando formados pelo *book-to-market*, investimento ou VPIN, os modelos não apresentaram desempenhos satisfatórios.

Por meio dos resultados apresentados pelo teste GRS, obtêm-se indícios de que um fator relacionado ao risco informacional, quando adicionado aos tradicionais modelos de fatores, seja apropriado na explicação dos retornos. De forma a analisar mais profundamente essa hipótese, procede-se à regressão dos fatores com base em Mohanram e Rajgopal (2009) e Fama e French (2015).

Tabela 9 Teste GRS para os modelos de 3, 4 e 5 fatores acrescidos do fator IMU

Combinações de fatores	Portfólios testados									
	12 portfólios: Tam. e BM.		12 portfólios: Tam. e Inv.		12 portfólios: Tam., Inv. e Luc.		10 portfólios: Tam. e VPIN		16 portfólios: Tam. e Luc.	
	GRS	Valor <i>p</i>	GRS	Valor <i>p</i>	GRS	Valor <i>p</i>	GRS	Valor <i>p</i>	GRS	Valor <i>p</i>
3 Fatores	1,7726	0,0498	1,6089	0,0853	1,0109	0,4372	1,8052	0,0570	0,8763	0,5970
3 Fatores + IMU	1,7336	0,0568	1,6116	0,0846	0,9900	0,4571	1,7529	0,0666	0,8844	0,5876
4 Fatores	1,7895	0,0470	1,6112	0,0847	1,0108	0,4372	1,8050	0,0571	0,8762	0,5973
4 Fatores + IMU	1,7554	0,0528	1,6121	0,0845	0,9913	0,4558	1,7538	0,0664	0,8854	0,5864
5 Fatores	1,7754	0,0493	1,6559	0,0734	1,0258	0,4232	1,7715	0,0630	0,8589	0,6176
5 Fatores + IMU	1,7459	0,0545	1,6699	0,0701	1,0076	0,4403	1,7318	0,0708	0,8677	0,6072

Tam.: Tamanho; BM: Book-to-Market; Inv.: Investimento; Luc.: Lucratividade.

A melhor combinação de fatores para cada conjunto de portfólios está destacada em negrito.

Fonte: Elaborada pelos autores.

4.4.2 Análise das relações entre os fatores de risco sistemático.

Esse procedimento, realizado por Mohanram e Rajgopal (2009) e Fama e French (2015), visa a testar se os interceptos das regressões são estatisticamente diferentes de 0. Um intercepto igual a 0 significaria que o fator não é precificado e que seu poder preditivo já estaria incorporado aos fatores existentes.

A Tabela 10 mostra as estimativas para as regressões dos 3 fatores de Fama e French (1993) e o fator IMU. O fator SMB apresenta intercepto estatisticamente diferente de 0, ou seja, os demais fatores não o incorporam aos seus retornos. Quando se adiciona o IMU, verifica-se substancial aumento do R² da regressão. Isso provavelmente se deve à forte correlação positiva de 0,56 entre os 2 elementos. No entanto, apesar desse aumento no poder de explicação da regressão, o fator SMB continua não sendo capturado

pelos demais.

Analisando o fator HML, este é o que apresenta maior valor *p* para seus interceptos. Os fatores MKT e IMU são os que mostram maior poder de explicação sobre o HML. Fama e French (2015) mostram que o HML se torna redundante para o modelo de 5 fatores. Os resultados da Tabela 10 sugerem que, mesmo para o modelo de 3 fatores, o HML não se apresenta como fator relevante, haja vista sua incorporação, em especial, pelo fator de mercado. Evidências utilizando o teste de GRS mostram que a exclusão do HML do modelo de 3 fatores, deixando somente o fator de mercado e o SMB, aumenta o poder de explicação do modelo, como será evidenciado na seção seguinte. Para o fator IMU não se rejeita a hipótese nula a respeito de seu intercepto. Este aparenta ser incorporado pelos demais fatores, em especial o SMB.

Tabela 10 Regressão dos três fatores de Fama e French (1993) e o IMU

Variável regredida	α	Variáveis independentes				R ²
		MKT	SMB	HML	IMU	
MKT	Coef.	-0,0008				0,3067
	Valor <i>p</i>	(0,148)				
	Coef.	-0,0008				0,3149
	Valor <i>p</i>	(0,188)				
SMB	Coef.	-0,001	-0,3418			0,1998
	Valor <i>p</i>	(0,003)	(0)			
	Coef.	-0,001	-0,2146			0,4023
	Valor <i>p</i>	(0,002)	(0)			
HML	Coef.	-0,004	0,3539	0,0827		0,1454
	Valor <i>p</i>	(0,427)	(0)	(0,127)		
	Coef.	-0,005	0,3623	-0,0505		0,1732
	Valor <i>p</i>	(0,334)	(0)	(0,413)		
IMU	Coef.	0	-0,0633	0,4004	0,105	0,3387
	Valor <i>p</i>	(0,314)	(0)	(0)	(0)	

Coef.: Coeficiente.

Fonte: Elaborada pelos autores.

Na análise das regressões pelos 4 fatores de Carhart (1997), ilustrados na Tabela 11, verifica-se que o fator UMD, apresentou os maiores valores *p* para seus interceptos, de 0,869 e 0,95, sendo este último relativo à adição do fator IMU. Isso significa que há fortes evidências de que os outros fatores capturam por completo o fator UMD.

Tabela 11 Regressão dos quatro fatores de Carhart (1997) e o IMU

Variável regredida	α	Variáveis independentes					R ²
		MKT	SMB	HML	UMD	IMU	
MKT	Coef.	0					0,3183
	Valor <i>p</i>	(0,159)					
	Coef.	0					0,3242
	Valor <i>p</i>	(0,195)					
SMB	Coef.	-0,001	-0,35				0,215
	Valor <i>p</i>	(0,004)	(0)				
	Coef.	-0,001	-0,2227				0,4197
	Valor <i>p</i>	(0,003)	(0)				
HML	Coef.	-0,0002	0,1976	-0,0201			0,3795
	Valor <i>p</i>	(0,556)	(0)	(0,667)			
	Coef.	-0,0003	0,2064	-0,1485			0,4059
	Valor <i>p</i>	(0,433)	(0)	(0)			
UMD	Coef.	0	-0,0807	-0,1099	-0,3781		0,3451
	Valor <i>p</i>	(0,869)	(0)	(0)	(0)		
	Coef.	0	-0,0736	-0,1547	-0,39		0,3507
	Valor <i>p</i>	(0,950)	(0,004)	(0)	(0)		
IMU	Coef.	0	-0,0557	0,4111	0,141	0,0933	0,3443
	Valor <i>p</i>	(0,320)	(0)	(0)	(0)	(0,020)	

Coef.: Coeficiente.

Fonte: Elaborada pelos autores.

Finalmente, em relação às regressões dos 2 fatores adicionados por Fama e French (2015), apresentados na Tabela 12, tem-se que o HML é, novamente, capturado pelos demais fatores. Os fatores RMW e CMA apresentaram interceptos próximos de 0 e com alto valor *p*.

Portanto, a partir das estimativas das regressões dos fatores apresentadas nas tabelas 10 a 12, verifica-se que os 2 principais fatores responsáveis pela explicação dos retornos das carteiras analisadas aparentam ser o MKT e o SMB, sendo os demais incorporados à medida que

outros fatores são adicionados às regressões. De forma a verificar qual combinação de fatores resulta no melhor modelo, realiza-se o procedimento estabelecido por Fama e French (2015) de análise dos resultados do teste GRS para diferentes arranjos de modelos.

Tabela 12 Regressão dos 5 fatores de Fama e French (2015) e o IMU

Variável regredida	α	Variáveis independentes						R ²
		MKT	SMB	HML	RMW	CMA	IMU	
MKT	Coef.	0,002						0,3213
	Valor <i>p</i>	(0,261)						
MKT	Coef.	0,001	-0,4199	0,5817	0,4188	0,0958	-0,2652	0,334
	Valor <i>p</i>	(0,326)	(0)	(0)	(0)	(0,168)	(0)	
SMB	Coef.	-0,001	-0,2722		0,0517	0,0253	0,4315	0,2739
	Valor <i>p</i>	(0,004)	(0)		(0,227)	(0,641)	(0)	
SMB	Coef.	-0,001	-0,1769		0,04	0,0413	0,289	0,5318
	Valor <i>p</i>	(0,003)	(0)		(0,296)	(0,145)	(0)	(0)
HML	Coef.	-0,0003	0,3266	0,0553		-0,7332	0,4262	0,4945
	Valor <i>p</i>	(0,463)	(0)	(0,227)		(0)	(0)	
HML	Coef.	-0,0003	0,3269	0,0535		-0,7327	0,4257	0,0049
	Valor <i>p</i>	(0,462)	(0)	(0,296)		(0)	(0)	(0,936)
RMW	Coef.	-0,0001	0,1551	0,0168	-0,4554		0,2105	0,3295
	Valor <i>p</i>	(0,593)	(0)	(0,641)	(0)		(0)	
RMW	Coef.	-0,0001	0,1446	0,0585	-0,4502		0,2203	-0,1126
	Valor <i>p</i>	(0,662)	(0)	(0,145)	(0)		(0)	(0,018)
CMA	Coef.	0,0002	0,0277	0,3392	0,313	0,2489		0,2959
	Valor <i>p</i>	(0,593)	(0,324)	(0)	(0)	(0)		
CMA	Coef.	0,0001	0,0389	0,2786	0,3076	0,2591	0,1483	0,3058
	Valor <i>p</i>	(0,678)	(0,168)	(0)	(0)	(0)	(0,004)	
IMU	Coef.	0,0002	-0,078	0,3715	0,002	-0,096	0,1075	0,3471
	Valor <i>p</i>	(0,364)	(0)	(0)	(0,935)	(0,018)	(0,004)	

Coef.: Coeficiente.

Fonte: Elaborada pelos autores.

4.4.3 Resultados para o teste GRS em diferentes combinações de modelos.

Para encontrar o modelo que melhor explica os retornos dos portfólios, Fama e French (2015) testam diferentes

arranjos de variáveis. Esse mesmo procedimento foi realizado com o objetivo de verificar qual combinação de fatores produz o melhor modelo para a amostra analisada (Tabela 13).

Tabela 13 Teste GRS para as variações dos modelos de 3, 4 e 5 fatores e o fator IMU

Combinações de fatores		Portfólios testados									
		12 portfólios: Tam. e BM		12 portfólios: Tam. e Inv.		12 portfólios: Tam., Inv. e Luc.		10 portfólios: Tam. e VPIN		16 portfólios: Tam. e Luc.	
		GRS	Valor p	Valor p	GRS	Valor p	GRS	Valor p	GRS	Valor p	GRS
3 Fatores + 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 Fatores + 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 Fatores + 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	

A coluna Modelos mostra quais combinações de fatores foram utilizadas como variáveis independentes na explicação dos retornos dos portfólios (variável dependente nas regressões) evidenciados nas demais colunas.

A melhor combinação de fatores para cada conjunto de portfólios está destacada em negrito.

Tam.: Tamanho; BM.: Book-to-Market; Inv.: Investimento; Luc.: Lucratividade.

Fonte: Elaborada pelos autores.

Analisando o conjunto de carteiras formadas pelo tamanho e *book-to-market* das ações, verifica-se que, dentre os 16 arranjos de fatores analisados, o que apresentou melhor performance foi aquele com os fatores MKT, SMB, UMD e IMU. Uma questão surge quando a inclusão do UMD melhora o desempenho dos modelos, tendo em vista que ele foi o que apresentou menor intercepto e com maiores valores *p* na Tabela 11. A explicação repousa na relação entre o UMD e o HML. Analisando a regressão do UMD com os demais fatores, constata-se que o HML apresenta o maior coeficiente absoluto dentre as variáveis, indicando, portanto, que esse seria o fator que capturaria melhor a variação dos retornos do UMD. Em resultados não reportados, as regressões desse fator com o MKT, SMB e, posteriormente, com a adição do IMU, verifica-se substancial diminuição do valor *p* do intercepto, o que sugere que esses fatores não conseguem capturar completamente o fator UMD, deixando margem para que tenha um papel na explicação dos retornos dos portfólios.

Também se constata que a adição do IMU melhora o desempenho de todos os modelos analisados. Novamente, deve-se buscar a explicação para essa evidência, já que, via regressões dos fatores, verificou-se que o IMU era

capturado pelo SMB e pelo HML. O SMB é o fator que tem maior poder de explicação sobre o IMU. Como o IMU tem o objetivo de captar a parte informacional das ações e a relação com o tamanho das empresas é uma consequência de como o mercado de capitais lida com o conteúdo informacional das empresas, é possível que o IMU seja capaz de explicar parte das variações dos retornos das carteiras não captadas pelo fator SMB – a parte relacionada ao risco informacional –, o que explicaria a melhora dos desempenhos dos modelos que incluem o fator IMU.

Em conclusão, verificou-se que as melhores combinações de fatores foram aquelas que excluíram o HML e incluíram o IMU. O modelo MKT, SMB e IMU foi o que apresentou melhor desempenho para 3 dos 5 conjuntos de carteiras. O modelo MKT, SMB, UMD e IMU apresentou performance superior para os portfólios formados pelo tamanho e *book-to-market*, porém, demonstrou também bom desempenho nos testes com os demais conjuntos, com exceção da combinação entre tamanho e lucratividade. Para esta última, a melhor combinação foi o modelo de 5 fatores, seguido do modelo de 4 fatores (MKT, SMB, RMW e CMA). A próxima seção trabalha com simulações dos retornos das carteiras

para verificar qual dessas combinações de fatores melhor explica os retornos na amostra analisada.

4.4.4 Análise dos modelos por meio do método *Bootstrap*.

Esta seção destaca o melhor modelo entre as seguintes combinações de fatores: MKT, SMB, IMU; MKT, SMB, UMD, IMU; MKT, SMB, HML, RMW, CMA; MKT, SMB, RMW, CMA. Optou-se por incluir os modelos de 5 fatores de Fama e French (2015) e sua versão restrita sem o fator

HML, com o objetivo de verificar, por meio de simulações, o efeito da exclusão do HML no desempenho do modelo. Para isso, utilizou-se o método *bootstrap* para a simulação dos retornos dos portfólios. Para a verificação da hipótese de que os interceptos não são estatisticamente diferentes de 0 foi utilizado o *Average F-test*. Por fim, de forma a determinar qual desses modelos apresentou melhor desempenho geral, utilizou-se o método de Fisher para combinação dos valores p apresentados pelos modelos, como mostrado por (7):

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

onde $\ln X_i$ é o logaritmo natural de cada valor p . Os resultados para as combinações dos valores p são apresentados na Tabela 14. Para cada conjunto de portfólios, um modelo superou os demais. Dessa forma,

fez-se o cálculo da combinação dos valores p das 62 simulações – relacionados às 62 carteiras dos 5 conjuntos formados.

Tabela 14 Resultado da combinação dos valores p por meio do método de Fisher

Combinações de fatores	Portfólios testados					Geral
	Tam. (3) e BM. (4)	Tam. (3) e Inv. (4)	Tam.(2), Inv. (3) e Luc. (2)	Tam. (3) e VPIN (4)	Tam. (4) e Luc. (4)	
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	$2,29 \times 10^{-7}$	$5,38 \times 10^{-6}$	$1,11 \times 10^{-6}$	$3,21 \times 10^{-5}$	$2,27 \times 10^{-6}$	$5,22 \times 10^{-24}$

A melhor combinação de fatores para cada conjunto de simulações de portfólios está destacada em negrito.

Tam.: Tamanho; BM.: Book-to-Market; Inv.: Investimento; Luc.: Lucratividade.

Fonte: Elaborada pelos autores.

O modelo MKT, SMB e IMU, apesar de ter apresentado melhores resultados para os conjuntos, como apresentado na Tabela 14, não se sustentou em relação aos demais quando aplicado aos dados simulados. A inclusão do fator de momento, resultando no modelo MKT, SMB, UMD e IMU, entretanto, conseguiu manter um bom desempenho nas simulações das carteiras, o que sugere que o fator UMD pode capturar variações dos retornos de forma geral na amostra analisada.

Finalmente, devido à evidência apresentada na Tabela

13, de que a inclusão do fator IMU no lugar do fator HML leva, de forma geral, a uma melhora no desempenho dos modelos, foi analisada a performance de um quinto modelo, não reportado na Tabela 14: MKT, SMB, RMW, CMA e IMU. O resultado da combinação dos valores p das simulações foi de $7,72 \times 10^{-24}$, ou seja, superior ao desempenho do modelo MKT, SMB, RMW e CMA de $5,22 \times 10^{-24}$, o que indica que a adição do fator relacionado ao risco informacional levou a melhora do modelo restrito de Fama e French (2015).

5. CONSIDERAÇÕES FINAIS

Este estudo buscou: (i) analisar a VPIN ou o nível de toxicidade das ações no mercado brasileiro; e (ii) verificar, via modelos de fatores propostos por Fama e French (1993, 2015) e Carhart (1997), se um fator de risco sistemático relacionado ao conteúdo informacional das

ações é precificado pelos investidores da BM&FBOVESPA.

Foi encontrado um VPIN médio de 0,4548, com desvio padrão igual a 0,2219 para o mercado brasileiro. Na análise relacionada aos segmentos de listagem das ações, foi verificado que o segmento N1 apresentou menor

VPIN, seguido do Tradicional, NM e N2. Esperava-se que ações do NM apresentassem menor VPIN, visto que o objetivo da segmentação da BM&FBOVESPA é levar maior transparência ao investidor, o que implicaria menor probabilidade de negociações privilegiadas. Os resultados sugerem que a probabilidade de negociações privilegiadas dos segmentos se relaciona com a quantidade de empresas e com as características das ações que os compõem, em especial, seu valor de mercado. Dessa forma, a hipótese estipulada nesta pesquisa, de que os segmentos teoricamente mais transparentes da BM&FBOVESPA apresentariam menor probabilidade de negociações privilegiadas, não pôde ser confirmada.

A hipótese de que existe uma correlação negativa entre o tamanho e a VPIN das ações foi corroborada. Os resultados deste estudo indicam que há uma correlação de $-0,3080$ entre o valor de mercado e a VPIN das empresas. A amostra analisada foi dividida em 3 grupos relacionados ao valor de mercado das companhias: pequeno, médio e grande. A VPIN média para esses grupos foi de $0,6364$, $0,3989$ e $0,3164$, respectivamente, indicando clara diminuição da VPIN na medida em que o tamanho das empresas cresce.

A última hipótese, referente ao papel de um fator relacionado ao risco informacional das ações, foi analisada por meio da construção do fator IMU. Acrescentou-se esse fator aos modelos de 3 e 5 fatores de Fama e French (1993, 2015) e 4 fatores de Carhart (1997) e utilizaram-se, como variáveis dependentes nas regressões, os retornos de 62 portfólios construídos com base no tamanho, *book-to-market*, lucratividade, investimento e VPIN das ações. Para todos os modelos, a adição do fator IMU aumentou o poder preditivo do SMB na amostra analisada.

De forma geral, a melhora no desempenho dos modelos com a inclusão do IMU foi verificada por meio do teste GRS. Para investigar mais detalhadamente essa afirmação, foram realizadas regressões entre os fatores. Os resultados indicam que todos os fatores, exceto o SMB e MKT são, em algum momento, capturados pelos demais. Fama e French (2015) atentam que, no contexto do modelo de 5 fatores, o fator HML se torna redundante. De forma a verificar quais desses fatores auxiliam na explicação dos retornos dos portfólios, o teste GRS foi aplicado em diferentes combinações de fatores.

Os resultados indicam que os modelos seguintes apresentaram melhor desempenho: MKT, SMB e IMU; MKT, SMB, UMD e IMU; MKT, SMB, HML, CMA e RMW; e MKT, SMB, CMA e RMW. Além de corroborar o resultado de Fama e French (2015), de que o fator HML é redundante para o modelo de 5 fatores, estendeu-se essa afirmação aos modelos de 3 e 4 fatores, tendo sido constatado, inclusive, que, quando presente, o HML

prejudica o desempenho dos modelos.

Para estender essas conclusões, realizou-se o procedimento de *bootstrap* dos retornos dos portfólios sendo eles regredidos com os modelos citados no parágrafo anterior. Posteriormente, aplicou-se o Average F-test. Os resultados para esse teste indicam que o modelo que melhor explica os retornos simulados é o MKT, SMB, RMW e CMA. A partir das evidências apresentadas anteriormente de que o fator IMU auxilia no desempenho dos modelos, procedeu-se à simulação dos retornos com o modelo MKT, SMB, RMW, CMA e IMU. O resultado encontrado foi que este último apresenta performance superior à dos demais modelos, o que fornece suporte para a hipótese central deste estudo de precificação do risco informacional no mercado acionário brasileiro.

A partir dos resultados das estimativas, entende-se que o fator IMU funciona como um complemento ao fator SMB – este último vital para o desempenho dos modelos – relacionado ao risco informacional das ações. A explicação se dá pela forma como ambos são construídos, já que as empresas menores (*small*) estão fortemente presentes no grupo *informed* e as grandes empresas (*big*) compõem o grupo *uninformed*. Caso esses 2 fatores fossem *proxies* um para o outro, o teste VIF apresentaria valor alto. Além disso, verificou-se que a correlação entre os 2, apesar de positiva, não é suficiente para que um fator incorpore por completo o outro, fato este evidenciado pela regressão do IMU com os demais fatores.

Concluindo, o fator relacionado ao risco informacional aparenta ter um papel significativo na explicação dos retornos dos portfólios criados. O fator de mercado e o SMB são os mais importantes no desempenho dos modelos, enquanto o HML não só é redundante como prejudicial. Os fatores adicionados por Fama e French (2015) ajudam a compor o modelo que melhor explica os retornos das 62 carteiras analisadas neste estudo.

A principal limitação desta pesquisa reside no período analisado. Enquanto trabalhos de precificação tendem a cobrir o maior período possível, não foi possível analisar um período maior que 2 anos. Essa limitação é compensada, em parte, pela utilização dos dados reais no cálculo da VPIN, procedimento este raro nos trabalhos de microestrutura de mercado pela não possibilidade de obtenção dos dados pela maior parte dos pesquisadores.

Como guia futuro das pesquisas nesta área, sugerem-se maiores análises da relação entre o fator SMB e IMU, objetivando entender a profundidade do relacionamento entre elas. Outra lacuna se refere à extração do conteúdo informacional das transações, ramo este liderado por Easley, López de Prado e O'Hara (2016) e que ainda necessita de muito avanço, por se tratar de uma variável de extrema complexidade e volatilidade.

REFERÊNCIAS

- Abad, D., & Rubia, A. (2005). Modelos de estimación de la probabilidad de negociación informada: una comparación metodológica en el mercado español. *Revista de Economía Financiera*, 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.
- Albarez, T., Lima, G., Lopes, A. & Valle, M. (2010). The relationship of asymmetric information in Brazilian public companies. *Review of Business*, 31, 3-21.
- Albarez, 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 intradiária 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). Recuperado de 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). Recuperado de 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). Recuperado de 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). Recuperado de https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2622782
- Fama, E., & French, K. (2017). Choosing factors (SSRN Working Paper). Recuperado de 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.

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

Endereço para correspondência:

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
E-mail: leonardosiq91@gmail.com