

**Universidade Federal de Pernambuco
Centro de Ciências Sociais Aplicadas
Departamento de Ciências Administrativas
Programa de Pós-Graduação em Administração - PROPAD**

Ricardo Cavalcante Galvão

**Finanças comportamentais: uma proposta de índice
para aferição do sentimento do investidor no
mercado de capitais brasileiro.**

**Recife
2016**

Ricardo Cavalcante Galvão

**Finanças comportamentais: uma proposta de índice
para aferição do sentimento do investidor no
mercado de capitais brasileiro.**

**Orientador: Profa. Dra. Joséte Florencio dos Santos
Coorientador: Prof. Dr. Pierre Lucena Raboni**

Tese de Doutorado apresentada como requisito complementar para obtenção do grau de Doutor em Administração, área de concentração em Finanças, do Programa de Pós-Graduação em Administração da Universidade Federal de Pernambuco.

**Recife
2016**

Catálogo na Fonte
Bibliotecária Ângela de Fátima Correia Simões, CRB4-773

G182f Galvão, Ricardo Cavalcante
Finanças comportamentais: uma proposta de índice para aferição do sentimento do investidor no mercado de capitais brasileiro / Ricardo Cavalcante Galvão. - 2016.
97folhas: il. 30 cm.

Orientadora: Prof.^a Dra. Joséte Florencio dos Santos e Coorientador Prof. Dr. Pierre Lucena Raboni
Tese (Doutorado em Administração) – Universidade Federal de Pernambuco. CCSA, 2016.
Inclui referência e apêndices.

1. Finanças. 2. Índices de mercado de ações. 3. Sociedades comerciais – Relações com investidores. I. Santos, Joséte Florencio dos (Orientadora). II. Raboni, Pierre Lucena (Coorientador). III. Título.

658 CDD (22. ed.) UFPE (CSA 2017 – 181)

Universidade Federal de Pernambuco
Centro de Ciências Sociais Aplicadas
Departamento de Ciências Administrativas
Programa de Pós-Graduação em Administração - PROPAD

Finanças comportamentais: uma proposta de índice para aferição do sentimento do investidor no mercado de capitais brasileiro.

Ricardo Cavalcante Galvão

Tese submetida ao corpo docente do Programa de Pós-Graduação em Administração da
Universidade Federal de Pernambuco e aprovada em 12 de dezembro de 2016.

Banca Examinadora:

Profa. Dra. Joséte Florencio dos Santos, UFPE (Orientadora)

Profa. Dra. Umbelina Cravo Teixeira Lagioia (Examinadora Interna)

Prof. Dr. André Magalhães, UFPE (Examinador Externo)

Prof. Dr. Pierre Lucena Raboni, UFPE (Examinador Externo)

Prof. Dr. Odilon Saturnino Silva, IFPB (Examinador Externo)

Dedico esta obra à santíssima trindade.

Agradecimentos

Agradeço a Deus por permitir e possibilitar a elaboração de uma tese e de chegar ao fim do curso de doutorado, além das inúmeras bênçãos já concedidas e proteções despendidas. Cada fase do curso foi acompanhada pela proteção divina a qual agradeço e pretendo honrar.

Agradeço a minha esposa Gabriella Galvão pelo apoio, suporte, dedicação e respeito ao longo desses tão atribulados anos, ao meu pai Gilvam George Galvão por ter concedido valorosos auxílios em períodos importantes.

Aos professores orientadores que guiaram a mim ao longo desses quase quatro anos, Pierre Lucena e Joséte Florencio dos Santos. Agradeço também à professora Umbelina Lagioia, a quem tenho gratidão e um profundo respeito por ter sido auxiliado sempre, e aos demais professores que me permitiram evoluir academicamente ao longo dos anos.

Aos docentes e colegas alunos do Propad-UFPE, que agem com o respeito, o companheirismo e a dedicação que nos faz sentir acolhidos e prontos a evoluir academicamente, notadamente os amigos Valéria e Odilon Saturnino, aos quais devo parte dessa conquista.

Aos amigos, parceiros e colegas, por terem dedicado atenção e bons momentos. À diretoria do CAA-UFPE por ter acreditado na qualificação do corpo de técnicos administrativos e ter concedido o tempo necessário ao desenvolvimento deste trabalho. Aos chefes Karina Carvalho e Henrique de Queiroz Chaves pela compreensão e apoio constantes.

"Vocês são a luz para o mundo. Não se pode esconder uma cidade construída sobre um monte. Ninguém acende uma lamparina para colocá-la debaixo de um cesto. Pelo contrário, ela é colocada no lugar próprio para que ilumine todos os que estão na casa. Assim também a luz de vocês deve brilhar para que os outros vejam as coisas boas que vocês fazem e louvem o Pai de vocês, que está no céu."

Mateus 5:14-16

Resumo

Abordando a temática finanças comportamentais, o presente trabalho teve por objetivo verificar o efeito no retorno das ações da BM&FBOVESPA quando considerado o sentimento do mercado norte-americano na propensão à geração de vieses dos investidores brasileiros no período de 2006 a 2015. Para atingir o objetivo proposto, foi criado um índice de sentimento mensal que teve por base os trabalhos de Baker e Wurgler (2006) e Yoshinaga (2009), mas com a inclusão de uma variável representativa do sentimento norte-americano, que representou o contágio de sentimento entre países. O índice foi obtido através da extração do primeiro fator de uma análise de componentes principais na qual fora observada a validade do índice norte-americano, a diferença entre os índices *market-to-book* das empresas pagadoras e não pagadoras de dividendos, o número e o retorno das ofertas públicas iniciais e o índice de confiança do consumidor, sendo apenas a última variável excluída por não compor o primeiro fator. A significância estatística do índice proposto foi verificada através da técnica de regressão com dados em painel, mais precisamente um painel de efeitos fixos composto pelas variáveis propostas no modelo de cinco fatores de Fama e French (2014), além do referido índice. Os resultados indicaram que o índice de sentimento proposto foi estatisticamente significativo ao nível de um por cento e com resultados melhores do que o mesmo índice sem a inclusão de variável representativa de contágio. Concluiu-se que há evidências no mercado de capitais brasileiro tanto de interferência dos aspectos comportamentais no retorno dos ativos quanto de contágio do sentimento do investidor norte-americano também no mercado brasileiro.

Palavras-chave: Finanças comportamentais. Índice de sentimento do investidor. Contágio de sentimento.

Abstract

Addressing behavioral finance, The objective of this study was to verify the effect on the return of BM&F BOVESPA stock market when considering the US investor sentiment in the propensity to generate biases of Brazilian investors in the period from 2006 to 2015. To reach the objective a monthly sentiment index was created based on Baker and Wurgler (2006) and Yoshinaga (2009) but with an extra variable representing the north american sentiment, which represented the contagion between countries. The index was obtained by extracting the first factor from a principal components analysis with these variables: North American sentiment index, the difference between market-to-book indices of paying and not paying dividend firms, the number and return of initial public offerings, and the consumer confidence index, being the former excluded because it did not make up the first factor. The statistical significance of the proposed index was verified using fixed effects panel, composed with the variables proposed in Fama and French (2014) the five-factor model and the index proposed. The results show that the index of sentiment was statistically significant at the level of one percent and obtained better results than the sentiment itself without the contagion variable. The variables included in the test were statistically significant except for profitability. The conclusion shows evidence of interference of behavioral aspects in the return of Brazilian stock market and also evidence of contagion of the US investor sentiment in the Brazilian stock market.

Keywords: Behavioral finance. Investor sentiment index. Sentiment contagion.

Lista de figuras

Figura 1 (2): Função de utilidade.....	25
Figura 2 (2): Função de valor proposta pela teoria da perspectiva.....	32
Figura 3 (2): Função de ponderação.....	34
Figura 4 (3): Composição das carteiras através do tamanho e do book-to-market.....	57
Figura 5 (3): Composição das carteiras através do tamanho e da lucratividade.....	58
Figura 6 (3): Composição das carteiras através do tamanho e do investimento.....	59
Figura 7 (4): Scree plot dos componentes da ACP.....	70

Lista de tabelas

Tabela 1(4): Estatística descritiva das variáveis e suas defasagens.....	68
Tabela 2(4): Cargas fatoriais da primeira análise de componentes principais.....	69
Tabela 3(4): Componentes da análise de componentes principais.....	71
Tabela 4(4): Cargas fatoriais da segunda análise de componentes principais.....	71
Tabela 5(4): Componentes da ACP excluindo a variável contágio.....	72
Tabela 6(4): Cargas fatoriais da análise de componentes principais sem a inclusão da variável contágio.....	73
Tabela 7(4): Estatística descritiva das carteiras Tamanho/Book-to-market.....	75
Tabela 8(4): Estatística descritiva das carteiras Tamanho/Lucratividade.....	75
Tabela 9(4): Estatística descritiva das carteiras Tamanho/Investimento.....	76
Tabela 10(4): Retornos médios das carteiras que compuseram os fatores.....	78
Tabela 11(4): Valores médios dos fatores HML, RMW, CMA e SMB.....	78
Tabela 12(4): Resultados dos testes GRS.....	79
Tabela 13(4): Diagnóstico de Painel.....	82
Tabela 14(4): Painel de efeitos fixos com carteiras tamanho e book-to-market.....	84
Tabela 15(4): Painel de efeitos fixos com carteiras tamanho e lucratividade.....	85
Tabela 16(4): Painel de dados empilhados com carteiras tamanho e investimento.....	86
Tabela 17(4): Escore z dos coeficientes angulares das variáveis de sentimento.....	88

SUMÁRIO

1 INTRODUÇÃO	13
1.1 PROBLEMA E PERGUNTA DE PESQUISA.....	15
1.2 OBJETIVOS.....	16
1.3 JUSTIFICATIVA.....	16
2 REFERENCIAL TEÓRICO	20
2.1 O MODELO CAPM E A HIPÓTESE DE MERCADOS EFICIENTES.....	20
2.2 MODELOS DE RACIONALIDADE DO INVESTIDOR.....	23
2.3 TEORIA DA PERSPECTIVA.....	29
2.3.1 A função de valor	32
2.3.2 A função de ponderação	33
2.3.3 Vieses comportamentais	34
2.4 ÍNDICES DE SENTIMENTO.....	36
3 PROCEDIMENTOS METODOLÓGICOS	43
3.1 DELINEAMENTO, POPULAÇÃO, AMOSTRA E VARIÁVEIS.....	43
3.1.1 Variáveis do índice de sentimento	44
3.1.2 Variáveis do modelo de apreçamento	47
3.2 COLETA E ORGANIZAÇÃO DOS DADOS.....	49
3.3 ÍNDICE DE AFERIÇÃO DE VIESES NA PRECIFICAÇÃO DE ATIVOS.....	52
3.3.1 Análise de componentes principais	52
3.4 MODELO DE AFERIÇÃO DA INFLUÊNCIA DO ÍNDICE NO MERCADO.....	55
3.4.1 Construção das carteiras	56
3.4.2 Construção dos fatores	59
3.4.3 Teste GRS e regressões com dados em painel	62
3.5 APRESENTAÇÃO DAS HIPÓTESES PROPOSTAS.....	65
4 RESULTADOS	67
4.1 ÍNDICE DE SENTIMENTO.....	67
4.1.1 Estatísticas descritivas das variáveis utilizadas nos índices de sentimento	68
4.1.2 Análise de componentes principais com todas as variáveis e suas defasagens	69
4.1.3 Análise de componentes principais com o indicador de contágio	70
4.1.4 Análise de componentes principais sem o indicador de contágio	72
4.2 AFERIÇÃO DA INFLUÊNCIA DO SENTIMENTO NO RETORNO DOS ATIVOS.....	74
4.2.1 Análise descritiva das carteiras	74
4.2.2 Análise descritiva dos fatores	76
4.2.3 Teste GRS	78
4.2.4 Aferição dos efeitos do sentimento no retorno dos ativos	80
4.2.4.1 Escolha do modelo de regressão.....	81
4.2.4.2 Resultados das regressões com as carteiras tamanho e <i>book-to-market</i>	82
4.2.4.3 Resultados das regressões com as carteiras tamanho e lucratividade.....	84
4.2.4.4 Resultados das regressões com as carteiras tamanho e investimento.....	86
4.2.3 Comparação entre os índices de sentimento	87
5 CONSIDERAÇÕES FINAIS	90
REFERÊNCIAS	93
APÊNDICE A – SÉRIE HISTÓRICA DO ÍNDICE DE SENTIMENTO	97

1 Introdução

Ao longo das décadas do século XX, muitos acadêmicos e investidores estudaram formas de conseguir retornos superiores à média do mercado. Até a década de 1950, havia opiniões divergentes a respeito da melhor forma de gerenciar o risco de um investimento.

Com a teoria proposta por Markowitz (1952) de que, através da diversificação, o investimento em um *portfolio* de ativos é capaz de reduzir o risco global e manter a média da rentabilidade, estudos se concentraram em alternativas de redução do risco considerado não diversificável, bem como a correta mensuração deste.

Apesar da diversificação reduzir o risco total, parte deste não poderia ser reduzido apenas com a montagem de uma carteira composta por vários ativos, uma vez que havia risco não-diversificável, ou risco sistemático. Como os modelos microeconômicos da época não consideravam os investimentos em condição de risco, havia a necessidade de construção de um modelo que considerasse tal condição.

Para contornar o problema descrito, foi desenvolvido o modelo conhecido como *Capital Asset Pricing Model (CAPM)* por Sharpe (1964), Lintner (1965), e Mossin (1966). Considerado referência na precificação de ativos, o modelo CAPM considera a relação risco/retorno linear, adicionando um coeficiente β (beta) a um determinado ativo a fim de ajustar o seu valor ao risco do mercado. Dessa forma, ativos considerados de maior risco naturalmente oferecem prêmios maiores aos seus compradores.

Considerando o suposto equilíbrio entre risco e retorno dos ativos cotados no mercado através da adoção do modelo CAPM, Fama (1970) sugeriu que cada ativo seria precificado de acordo com o seu risco e retorno de forma a manter o mercado em equilíbrio. Surgiu então a Hipótese de Mercado Eficiente (HME) como sendo aquela na qual os preços dos ativos refletem toda a informação disponível em determinado momento. Assim, novas informações a respeito de cada ativo seriam rapidamente assimiladas pelo mercado, eliminando a possibilidade de ganhos extraordinários por parte do investidor.

Com o passar dos anos, críticas teóricas e evidências empíricas sugeriram que tanto o modelo CAPM quanto a HME não se aplicavam de maneira conclusiva no mercado. O trabalho conhecido como Roll's critique, proposto por Roll (1977), após alguns anos, repercutiu na academia por questionar os fundamentos e pressupostos do modelo CAPM.

Estudos empíricos indicaram a presença de fatores distintos dos coeficientes de risco propostos no modelo de Sharpe (1964), Lintner (1965), e Mossin (1966), contrariando o modelo e a HME. Tais variáveis foram então classificadas como anomalias de mercado. Dentre as principais anomalias observadas, houve aquelas relacionadas a fatores temporais, como diferenças de rentabilidade média entre a sexta-feira e a segunda-feira, outras relacionadas a fatores relacionados ao valor das empresas, como a relação entre o valor contábil do patrimônio líquido e o valor de mercado, além do tamanho da organização.

Os estudos envolvendo a contestação da suposição de eficiência do mercado podem ser divididos em dois grupos. O primeiro trata dos vieses cognitivos do ser humano e sua consequente influência nos mercados (KAHNEMAN; TVERSKY, 1979). O segundo trata de evidências de ineficiência encontradas no mercado de capitais ao longo dos anos (FAMA; FRENCH, 2014).

Dada a quantidade de evidências contestando a eficiência de mercado e a validade do modelo CAPM, parte das então chamadas anomalias de mercado passaram a fazer parte de um novo ramo das finanças, chamado finanças comportamentais.

Conforme descreve Lobão (2010, p. 122)

[...] as finanças comportamentais propõem a utilização de pressupostos mais realistas acerca dos investidores. Na verdade, os investidores não são seres totalmente racionais mas isso não significa que cometam erros flagrantes com frequência. Os investidores não podem saber tudo sobre todos os assuntos, têm diferentes graus de entendimento em relação a diferentes matérias e, por vezes, têm comportamentos altruístas. Têm capacidades cognitivas limitadas, o que os impede de processar toda a informação disponível, são influenciados pelas emoções e pelo contexto social e cultural em que são chamados a decidir. No fundo, o agente económico típico nas finanças comportamentais é como qualquer um de nós.

Nos últimos anos, foram estudados os mais diversos efeitos sobre o mercado de ações nos mais variados lugares do mundo, desde o efeito do ano novo lunar no mercado chinês (YUAN; GUPTA, 2014), até o viés comportamental dos mercados africanos (ALAGIDEDE, 2011), passando pelo mercado inglês (NICHOL; DOWLING, 2014), alemão (LOBE; RIEKS, 2011) e brasileiro (YOSHINAGA, 2009). Até as diferenças de comportamento entre os países considerados individualistas e coletivistas foram observadas no estudo de Eun, Wang e Chiao (2013), sugerindo que a cultura seria um fator importante, porém negligenciado na literatura do campo das finanças.

Outros estudos têm sido direcionados à proposição de índices de sentimento nos mercados, que seriam capazes de aferir graus de otimismo ou pessimismo perante os investidores e identificar se estes estariam tomando decisões fora dos padrões de

racionalidade. Diferentes variáveis foram propostas na tentativa de identificar o sentimento, desde *surveys* (BROWN; CLIFF, 2005), *proxies* isoladas (LEMMON; PORTNIAGUINA, 2006; SCHMELING, 2009), ou a partir de um conjunto de *proxies* (BAKER; WURGLER, 2006; BAKER; WURGLER; YUAN, 2012; YOSHINAGA, 2009).

Dada a necessidade de um diferente entendimento acerca do comportamento do investidor, substituindo o pressuposto de indivíduo totalmente racional anteriormente aceito, cabe o estudo das alterações no valor fundamental dos ativos em virtude das limitações cognitivas dos indivíduos, podendo este resultar em possibilidades tanto de ganho extraordinário quanto de melhoria na alocação de capital na economia.

1.1 Problema e pergunta de pesquisa

A teoria clássica de finanças toma como um de seus pressupostos a racionalidade do investidor, que analisaria as probabilidades de ocorrência dos cenários e suas consequências em termos de retorno sobre o investimento, resultando na escolha eficiente de uma carteira de ativos.

As evidências empíricas, no entanto, apresentaram resultados diferentes dos esperados. Foi verificado que os investidores apresentavam comportamentos distintos daqueles esperados, o que possibilitaria a obtenção de retornos acima da média com a manutenção do risco, os quais foram chamados de anomalias.

O erros de avaliação, conforme proposto por Khaneman e Tversky (1979) e apresentados no referencial teórico, são oriundos de vieses comportamentais do ser humano, que mensura as probabilidades de ocorrência de eventos de maneira distinta do proposto no modelo de racionalidade utilitária de Von Neumann e Morgenstern (1953), que determina o retorno provável de um ativo através da multiplicação das probabilidades pelos retornos possíveis, considerado como premissa dos modelos utilizados em finanças. Esses vieses causam alterações na precificação dos ativos, levando à ineficiência do mercado e consequente má alocação de recursos.

Outro aspecto relevante envolve a não observância de contágio entre sentimentos de diferentes países. Os trabalhos recentes envolvendo sentimento do investidor têm apresentado indícios de contágio entre países e, dada a integração entre os mercados financeiros e o fluxo constante de capitais entre países, pode-se supor a existência de contágio entre os sentimentos dos investidores também no Brasil. A esse respeito, Baker, Wurgler e Yuan

(2012) consideraram a ideia de replicar o índice de sentimento em diferentes mercados e sugerir um índice global. Foi então verificada a existência de correlação entre os índices nacionais, corroborando a ideia de contágio de sentimento nos mercados. A esse respeito, afirmaram que “O efeito do sentimento global sugere um mecanismo novo: sentimento pode ser contagioso” (BAKER; WURGLER; YUAN, 2012, p. 285, tradução nossa).

O problema de pesquisa desta tese envolve os erros de precificação de ativos causados pela influência de vieses oriundos de sentimentos como o otimismo ou o pessimismo por parte dos investidores e o possível contágio deste sentimento entre diferentes mercados.

Dada a problemática acima mencionada, tomando como base os trabalhos de Yoshinaga (2009) e Baker e Wurgler (2006) e adicionando variável capaz de captar o contágio entre mercados, este trabalho teve por finalidade oferecer resposta a seguinte pergunta de pesquisa: **Qual o efeito sobre o retorno das ações da BM&FBOVESPA, quando considerado o efeito contágio do sentimento do mercado norte-americano, no comportamento do investidor no mercado brasileiro?**

A partir da pergunta apresentada, pretendeu-se contribuir com o desenvolvimento de um índice de sentimento para o mercado de capitais brasileiro que tivesse em sua composição uma variável capaz de representar o sentimento de outros mercados, bem como as demais variáveis comuns aos estudos publicados.

1.2 Objetivos

O objetivo geral deste trabalho foi verificar o efeito no retorno das ações da BM&FBOVESPA quando considerado o sentimento do mercado norte-americano na propensão à geração de vieses dos investidores brasileiros no período de 2006 a 2015.

Para que seja possível atingir o objetivo proposto, foram estabelecidos os seguintes objetivos específicos:

- 1 Mensurar a propensão à geração de vieses na precificação dos ativos por parte dos investidores brasileiros por meio da *proxy* índice de sentimento brasileiro – ISBR - e considerando o sentimento norte-americano proposto por Baker e Wurgler (2006);
- 2 Analisar a influência do ISBR no retorno das ações negociadas na BM&FBOVESPA no período entre 2006 e 2015;

1.3 Justificativa

A justificativa deste estudo se apresenta do ponto de vista teórico e prático. Do ponto de vista teórico, o estudo se justifica uma vez que as evidências de que os modelos clássicos de finanças não funcionam em sua plenitude sugerem que um ou mais pressupostos aceitos pela teoria clássica de finanças não correspondem à realidade do mercado, conforme apresenta Shiller (2003) em uma retrospectiva sobre a evolução do conhecimento sobre finanças a partir da hipótese de eficiência dos mercados.

Evidências da influência do sentimento do investidor e do contágio do sentimento entre mercados trazem a necessidade de esclarecer tais ocorrências. Dos estudos focados na temática sentimento, Baker e Wurgler (2006) utilizaram o modelo de três fatores de Fama e French (1993) e acrescentaram uma variável de sentimento. Yoshinaga (2009), por sua vez, utilizou o modelo CAPM acrescido das variáveis tamanho (BANZ, 1981), alavancagem (BHANDARI, 1988), book-to-market (FAMA e FRENCH, 1992) e momentum (CARHART, 1997), adicionando, por fim, a variável sentimento. Kurov (2010) utilizou o modelo CAPM adicionado de duas medidas de *spread* e o índice de sentimento. Para aferir o contágio, Baker, Wurgler e Yuan (2012) utilizaram um índice internacional e testaram através do modelo de três fatores mais o crescimento nas vendas.

Pode-se observar que não há um padrão consolidado de escolha de variáveis que compõem os modelos de teste de variáveis representativas do sentimento, sendo comum a todos apenas o CAPM e sendo particular a cada um as variáveis adicionais cujo propósito é melhorar o poder de explicação do retorno, aumentando a precisão do efeito do sentimento em cada teste. Quanto ao método, alguns pesquisadores optaram por agrupar as ações em carteiras enquanto outros optaram por comparar o retorno de cada ativo individualmente.

Neste trabalho, optou-se pelo modelo de cinco fatores proposto por Fama e French (2015) adicionado da variável sentimento. Segundo os autores, o desempenho do modelo de cinco fatores é superior ao do modelo de três fatores.

A busca por evidências de efeitos locais de fenômenos observados no exterior, como o contágio entre sentimentos de diferentes países, é um aspecto não adotado por outros estudos envolvendo o mercado de capitais brasileiro e que justifica o desenvolvimento deste trabalho.

Do ponto de vista prático, torna-se válida a busca pelo desenvolvimento e proposição de novos fatores ou variáveis que melhorem os modelos de precificação existentes. Encontrar fatores que alteram o preço dos ativos contribui para o aumento da eficiência do mercado.

Esta tese está dividida em cinco seções, sendo a primeira a introdução. A segunda seção, referencial teórico, está dividida em quatro subseções. A primeira trata do CAPM e da Hipótese de Mercados Eficientes, apresentando seus modelos e seus pressupostos. O intuito deste capítulo é apresentar a teoria clássica de finanças e seus pressupostos, fornecendo a base teórica sob a qual foi estruturado o problema de pesquisa.

A segunda subseção do referencial teórico apresenta os modelos de racionalidade do investidor, notadamente as obras de Bernoulli (1954) e Von Neumann e Morgenstern (1953). Tais obras fundamentaram os pressupostos de racionalidade adotados nos modelos clássicos de finanças, sendo a racionalidade um modelo normativo no qual se espera do ser humano determinado padrão de comportamento. Dado que os comportamentos humanos sugeridos em tais estudos serviram como pressupostos do modelo clássico de finanças, faz-se necessário apresentar tais modelos para fundamentar a sua crítica, a ser apresentada no capítulo seguinte.

A terceira subseção trata da Teoria da Perspectiva, proposta por Kahneman e Tversky (1979), que difere das teorias anteriores por apresentar um modelo empírico de racionalidade baseada no comportamento observado de indivíduos. O capítulo apresenta ainda, em subseções, a função de valor e a função de ponderação, ambas pertencentes à Teoria da Perspectiva e os vieses comportamentais observados na literatura, que são caracterizados por padrões de comportamento que levam indivíduos a abandonar o padrão de racionalidade proposto no modelo clássico de finanças. A quarta subseção apresenta os índices de sentimento criados, além de seus pressupostos teóricos e estratégias metodológicas adotadas pelos respectivos autores.

A terceira seção deste projeto trata dos procedimentos metodológicos escolhidos com o intuito de responder à pergunta de pesquisa e chegar aos objetivos propostos inicialmente. A seção está dividida em cinco subseções. A primeira subseção dos procedimentos metodológicos apresenta as variáveis escolhidas, tanto na construção da variável representativa do índice de procedimento quanto no modelo a ser utilizado como teste da significância da variável proposta.

A segunda subseção dos procedimentos metodológicos descreve a coleta e a organização dos dados que compuseram a amostra a ser utilizada no trabalho. A terceira subseção apresenta o índice de sentimento proposto.

A quarta subseção apresenta o modelo econométrico a ser utilizado no teste da variável proposta, que envolve o uso de regressão com dados em painel. Nesta foram

apresentados os testes validadores do modelo e a estrutura do modelo final de teste da variável, além das implicações dos possíveis resultados encontrados. Na quinta subseção, são apresentadas as hipóteses propostas.

2 Referencial Teórico

Nesta seção são apresentados os aspectos teóricos utilizados na fundamentação da tese, que, conforme apresentado na subseção 1.2, tem por objetivo construir um índice capaz de captar o sentimento do investidor brasileiro.

Para a devida fundamentação teórica, esta seção está dividida em quatro subseções, sendo a primeira relacionada às finanças clássicas, mais notadamente o modelo CAPM e a Hipótese de Mercados Eficientes – HME.

A segunda subseção trata dos modelos de racionalidade do investidor, sendo iniciada pela racionalidade proposta no modelo clássico de finanças, seguida pelas evidências iniciais de que o comportamento observado nos seres humanos não segue os padrões propostos.

Em seguida, na terceira subseção, é apresentada a teoria da perspectiva e suas evidências a partir de observações dos padrões de comportamento dos indivíduos. Por fim, na quarta subseção, são apresentados os índices de sentimento utilizados em estudos anteriores.

2.1 O modelo CAPM e a Hipótese de Mercados Eficientes

O *Capital Asset Pricing Model* proposto por Sharpe (1964) considera uma relação linear entre o risco e o retorno esperado de um determinado ativo ou carteira, representada através de uma taxa de desconto que considera a taxa livre de risco do mercado e um prêmio de risco do ativo ou carteira em questão. A representação do modelo se dá através da equação:

$$E(R_i) = R_f + \beta (E(R_m) - R_f) \quad (1)$$

Onde $E(R_i)$ significa a expectativa de retorno do ativo i , R_f significa a taxa livre de risco, β o coeficiente linear do modelo e $E(R_m)$ o retorno esperado do mercado.

Através do CAPM seria possível estabelecer o retorno esperado de qualquer ativo a partir da média e da variância dos retornos do respectivo ativo e do mercado. Os pressupostos do modelo, no entanto, foram motivo de controvérsia dada a sua dificuldade de aplicação ou atendimento por parte dos mercados.

Os pressupostos supõem que os investidores procuram maximizar o seu resultado econômico, são racionais e avessos a risco, possuem ativos diversificados, não exercem influência nos preços, podem emprestar e tomar emprestado quantidades ilimitadas de recursos pela taxa de juros livre de risco, não há custos de transação em suas operações de compra e venda de ativos, lidam com ativos divisíveis e líquidos, possuem expectativas homogêneas e é assumido que toda a informação está disponível ao mesmo tempo para todos os investidores.

Observa-se, no entanto, que alguns pressupostos, conforme apontou Roll (1976), são improváveis de se verificar em mercados. O quinto pressuposto postula que qualquer investidor pode contrair empréstimos pela mesma taxa de juros do governo. Este pressuposto é implausível porque o risco de *default* do governo é virtualmente nulo, uma vez que este pode imprimir moeda.

O oitavo pressuposto diz que todos os investidores possuem a mesma expectativa em relação aos resultados futuros da organização. Isto significa que é esperado que todos os indivíduos que investem no mercado acreditem no mesmo desempenho futuro de cada uma das empresas, sendo tal fato improvável. A mesma informação disponível a todos os investidores, nono pressuposto, implica em que nenhum indivíduo possui informação privilegiada e, principalmente, que nenhum investidor está desinformado.

Tais pressupostos são de difícil atendimento, sendo a viabilidade do modelo contestada não pela construção lógica, mas pela dificuldade de observação em um ambiente real.

A partir do modelo CAPM, Fama (1970) postulou que, uma vez atendidos os pressupostos por ele determinados, seria de se esperar que os investidores levassem os preços do mercado a um equilíbrio, impedindo que alguém obtivesse retornos sistematicamente acima da média do mercado por vários períodos mantendo o mesmo nível de risco. A essa ideia foi dada o nome de Hipótese de Mercados Eficientes (HME).

Seguindo a Hipótese de Mercados Eficientes, em um mercado em condições de equilíbrio, não seria possível auferir ganhos acima da média de mercado, uma vez que o mercado reagiria imediatamente a novas informações disponíveis, não permitindo ao investidor aproveitar os momentos de sub ou sobre valorização. Pelo modelo, retornos anormais, que seriam aqueles fora da linha traçada pela linha de retorno esperado do CAPM, retornariam o valor nulo [zero].

Ainda segundo Fama (1970), para que a HME funcione, algumas condições devem ser satisfeitas:

1. Não devem ser considerados os custos de transação;
2. Toda informação está disponível e sem custos aos participantes;
3. Todos concordam a respeito das implicações das informações atuais nos preços atuais, bem como daquelas ocasionadoras de alterações nas projeções futuras de cada ação;

A HME foi dividida em três níveis; fraca, semiforte e forte. Na sua forma fraca, não seriam possíveis ganhos anormais através da utilização de dados passados para previsão de resultados futuros, notadamente a análise técnica de ações. Na sua forma semiforte, não seriam possíveis ganhos acima da média de mercado para determinado nível de risco nem com os preços históricos nem com informações públicas. Na sua forma forte, não seria possível auferir ganhos extraordinários nem com informação privilegiada.

Conforme pode ser observado, as duas teorias oferecem simplificações que tornam possíveis avaliações objetivas sobre os ativos financeiros disponíveis no mercado. Se for levado em consideração o contexto da época de publicação do CAPM, tem-se ainda a limitada capacidade computacional da época, que impedia análises complexas envolvendo extensos e repetitivos cálculos.

Com o passar dos anos, críticas teóricas e evidências empíricas sugeriram que tanto o modelo CAPM quanto a HME não se aplicavam de maneira conclusiva no mercado. O trabalho conhecido como Roll's critique, proposto por Roll (1976), após alguns anos, repercutiu na academia por questionar os fundamentos e pressupostos do modelo CAPM.

Roll (1976) afirma que não seria possível testar a teoria proposta por Sharpe (1964), pois, em primeiro lugar, a única hipótese falseável é de que a carteira de mercado é de média-variância, o que limitaria a confirmar se a relação risco/retorno é linear. Segundo, o coeficiente de linearidade (β) não pode ser testado de maneira independente porque decorre de uma suposta eficiência do mercado. Terceiro, havendo infinitas combinações de carteiras, todas as relações entre risco e retorno seriam obrigatoriamente lineares. Outro grande problema estaria no fato de que não seria possível encontrar uma *proxy* que se aproximasse da carteira do mercado, pois seria necessário incluir todos os ativos disponíveis ao investidor, desde ações até imóveis, na composição da sua carteira.

Dadas as restrições impostas pelos pressupostos do modelo, vários comportamentos divergentes do proposto pela teoria clássica foram encontrados ao longo dos anos. A esses

padrões de comportamento foram dados os nomes de anomalias. Na revisão da literatura procedida por Carmona (2007), foram verificadas anomalias de calendário, fundamentais e técnicas. As anomalias de calendário são aquelas que envolvem características relacionadas ao horizonte temporal, incluindo os efeitos de determinado mês no retorno fora do esperado. Além das anomalias de calendário, foram identificados padrões divergentes do proposto pela teoria clássica em relação aos fundamentos das empresas, chamadas anomalias de fundamento ou fundamentais. Fama (1992) afirma que resultados de testes revelaram que variáveis como *Dividend Yield*, Relação Preço/Lucro, tamanho e estrutura de capital foram capazes de alterar o retorno de carteiras em relação ao previsto sem alteração no risco envolvido. Lucena e Pinto (2008) encontraram evidências das anomalias *book-to-market* e tamanho da firma no mercado brasileiro.

Em relação às anomalias técnicas, foram encontrados os efeitos *momentum* e *overreaction*. O efeito *momentum* consiste no padrão de manutenção dos retornos no curto prazo, ou seja, ações que obtiveram altos retornos tendem a manter esses retornos no curto prazo (BHOOTRA; HUR, 2013; HOU; MCKNIGHT, 2004; HUR; SINGH, 2014; JIANG; LI; LI, 2012; KIM et al., 2014; MENKHOFF et al., 2012). O efeito *overreaction* consiste na ideia de que as empresas que obtiveram retornos acima da média tendem a apresentar retorno inferior à média de mercado no médio prazo (DANIEL; HIRSHLEIFER; SUBRAHMANYAM, 1998; LOBE; RIEKS, 2011)(HUR; SINGH, 2014).

Dos modelos posteriores ao CAPM, destacam-se aqueles que agregam novas variáveis ao primeiro ao invés da sua total substituição. Fama e French (1992) rerepresentaram o CAPM com a inclusão de anomalias constatadas, tais como o efeito tamanho da empresa e a relação entre valor patrimonial e valor de mercado e posteriormente com a inclusão da lucratividade e do fator de investimento (FAMA; FRENCH, 2015). Também evidenciaram anomalias e propostas de alteração do modelo original Nichol e Dowling (2014), Israel e Moskowitz (2013) e Chou, Hou e Ko (2012).

Na subseção a seguir são apresentados os modelos de racionalidade do investidor que serviram de base para a fundamentação teórica que justificará a inclusão de um índice de sentimento no modelo de cinco fatores de Fama e French (2015).

2.2 Modelos de racionalidade do investidor

Uma classe de contestações do modelo clássico de finanças está na pressuposição de racionalidade utilitária do investidor. Os pressupostos de racionalidade apresentados se baseiam na obra de Von Neumann e Morgenstern (1953), cujos autores afirmaram que o tema abordado, racionalidade na tomada de decisão, ainda estava muito limitado, sendo o respectivo trabalho de caráter heurístico na transição entre modelos não matemáticos para modelos mais robustos e gerais. Porém, dada a possibilidade de simplificação de modelos a partir da abordagem proposta, muitos dos estudos posteriores não dedicaram a devida atenção a esta ressalva, passando apenas a considerar os investidores como racionais.

A ideia de racionalidade na tomada de decisão fora inicialmente baseada na obra de Bernoulli em 1738. Já há menção a concepções anteriores nas quais se tem a soma dos resultados multiplicados pela probabilidade de ocorrência destes como forma de mensurar o valor esperado de uma determinada operação probabilística. Porém, Bernoulli (1954, p. 24) afirmou que o valor de um ativo não deveria estar baseado no seu respectivo preço, mas na utilidade que este bem traz. Tal afirmação se baseia no conceito de que o preço é único para todos os indivíduos, porém, a utilidade depende das circunstâncias particulares de cada um.

Se a utilidade de cada expectativa de lucro possível é multiplicada pelo número de maneiras pelas quais isto pode ocorrer e se depois for dividida a soma desses produtos pelo número total de possíveis casos, uma utilidade média [expectativa moral]¹ será obtida e o lucro que corresponde a esta utilidade será igual ao valor do risco em questão. (BERNOULLI, 1954, p. 24, tradução nossa)

Percebe-se aqui que a utilidade dos resultados possíveis é o que leva o indivíduo a assumir riscos. O autor segue exemplificando que uma determinada quantia de dinheiro se faz muito mais útil àquele que dispõe de poucos recursos em comparação com um indivíduo suficientemente abastado, o que é um ponto chave na discussão sobre racionalidade na tomada de decisão, inclusive em teorias futuras. Bernoulli (1954, p. 25, tradução nossa) afirmou que a “utilidade resultante de qualquer pequeno acréscimo na riqueza será inversamente proporcional à quantidade de bens previamente possuídos”.

O conceito de que um indivíduo com menor quantidade de ativos tenderá a dar mais valor a uma determinada quantia em comparação com um indivíduo que dispõe de grande fortuna foi chave na discussão sobre a racionalidade. O aspecto decrescente da utilidade marginal alterou a escala da função, sendo a escala logarítmica mais adequada segundo o autor.

¹Bernoulli (1954) utilizou o termo valor moral da riqueza para se referir à utilidade gerada por uma determinada quantidade de dinheiro.

A Figura 1(2) apresenta a representação gráfica da função de utilidade proposta na qual o eixo das ordenadas representa a utilidade do capital e o eixo das abcissas representa a riqueza representada através do patrimônio do indivíduo. A curva BGLS representa a função de utilidade proposta e o segmento AB representaria o valor monetário dos bens do indivíduo inicialmente.

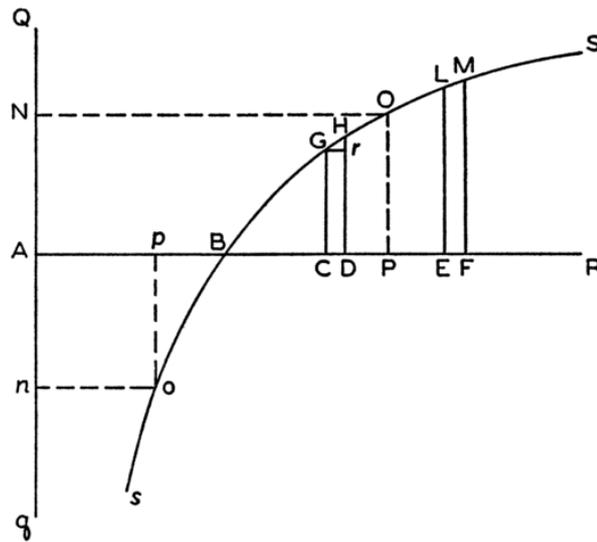


Figura 1 (2): Função de utilidade

Fonte: Bernoulli (1954, p. 26)

Seguindo a hipótese anteriormente aceita, para encontrar um valor de utilidade hipotético PO, deveriam ser somadas as probabilidades de ocorrência dos eventos CG, DH, EL e FM com a devida ponderação pela probabilidade de ocorrência de tais eventos, conforme fórmula abaixo adaptada de Bernoulli (1954, p. 26)

$$PO = \frac{\rho_1 CG + \rho_2 DH + \rho_3 EL + \rho_4 FM}{\rho_1 + \rho_2 + \rho_3 + \rho_4} \quad (2)$$

Onde ρ representa o número de possibilidades de ocorrência do evento dentro do espaço amostral. Analogamente, seria possível encontrar o valor monetário através da equação

$$PO = \frac{\rho_1 BC + \rho_2 BD + \rho_3 BE + \rho_4 BF}{\rho_1 + \rho_2 + \rho_3 + \rho_4} \quad (3)$$

Ao transformar a equação para a escala logarítmica, tem-se (BERNOULLI, 1954, p. 26):

$$AP = (AC^{\rho_1} \cdot AD^{\rho_2} \cdot AE^{\rho_3} \cdot AF^{\rho_4})^{\frac{1}{\rho_1 + \rho_2 + \rho_3 + \rho_4}} \quad (4)$$

Dada a natureza da curva de utilidade, nenhum jogo de azar valeria a pena aos jogadores uma vez que as perdas seriam mais sentidas do que os ganhos. A opção por empreender ou não poderia ser efetivamente calculada de acordo com o capital do indivíduo, limitando a proporção de comprometimento de sua riqueza nas situações de perda.

A proposição também permite a verificação da viabilidade de contratação de seguros. Novamente, indivíduos mais abastados teriam uma menor parcela do próprio capital comprometida com um ativo, sendo a contratação de seguro menos vantajosa do que para o indivíduo cujo valor do bem corresponderia a grande parte do seu patrimônio.

Entende-se aqui uma tentativa de estabelecimento de um modelo de caráter normativo sobre como deveria ser a tomada de decisão maximizadora da utilidade em condições de risco. Não se teve como objetivo postular sobre a racionalidade do ser humano ou elaborar uma teoria capaz de estabelecer um comportamento tido como racional. Von Neumann e Morgenstern (1953), por sua vez, tentaram apresentar um modelo matemático capaz de definir um comportamento racional. Acerca do objetivo, afirmaram:

Queremos encontrar os princípios matematicamente completos que definem ‘comportamento racional’ para os participantes de uma economia social, e que dela retire as características gerais desse comportamento. E, enquanto os princípios devem ser perfeitamente gerais - ou seja, válidos em todas as situações - podemos ficar satisfeitos se pudermos encontrar soluções, para o momento, apenas em alguns casos específicos. (VON NEUMANN e MORGENSTERN, 1953, p.31, tradução nossa)

Acerca do conceito de utilidade, há uma menção às tentativas anteriores de mensuração numérica da utilidade, consideradas ingênuas. Para os autores, ficou claro que qualquer mensuração será baseada, em última instância, em uma sensação. Para o caso da utilidade, a sensação imediata de preferência fornece a base para a sua mensuração.

Tal qual uma escala ordinal, a mensuração de utilidade apresentada serviria apenas para comparação entre duas opções disponíveis a um mesmo indivíduo, o que não permitiria a comparação entre as preferências de duas pessoas.

Há uma consideração importante a ser feita acerca da opinião dos autores quanto à natureza desta relação, que é a comparação com as ciências exatas, sugerindo que a subjetividade do tema remete mais a fase inicial do seu desenvolvimento do que a impossibilidade de quantificação do tema.²

²Um aspecto interessante da epistemologia empregada é a ideia de que o domínio da função utilidade contém operações precisas do ponto de vista matemático de maneira análoga às ciências exatas.

Para contornar a subjetividade da escala das preferências, os autores sugerem uma forma específica para quantificar as diferenças observadas, transformando a escala de ordinal para intervalar, que é a combinação de proporções entre as alternativas. Para encontrar uma correspondência entre números e utilidade foi apresentada a seguinte equação:

$$u \rightarrow \rho = v(u) \quad (5)$$

Sendo u a utilidade e $v(u)$ um valor correspondente a esta. Desta forma,

$$u > v \text{ implica em } v(u) > v(v)$$

$$v(\alpha u + (1 - \alpha)v) = \alpha v(u) + (1 - \alpha)v(v) \quad (6)$$

Sendo α a probabilidade de ocorrência dos eventos.

Para que as relações fossem consideradas válidas, foram propostos alguns axiomas (VON NEUMANN e MORGENSTERN, 1953, p. 24). O primeiro axioma é a completude, o qual afirma que, dadas duas opções, apenas três situações são possíveis:

$$u = v, u > v, u < v$$

De maneira intuitiva, dadas duas opções, pode haver preferência pela primeira, pela segunda ou indiferença quanto à preferência do indivíduo por uma destas.

O segundo axioma postula que, se u é preferível a v e v é preferível a w , então u será preferível a w .

$$u > v \text{ e } v > w \text{ implica em } u > w$$

O terceiro axioma leva em consideração três variáveis em uma hierarquia de preferências tal que $u > w > v$ implica na existência das seguintes possibilidades:

$$\alpha u + (1 - \alpha)v < w \quad (7)$$

Ou seja, dados três resultados possíveis para um evento com as preferências apresentadas acima, seria possível obter uma combinação de u e de v de modo que esta combinação viesse a ser preferível a w e *vice-versa*.

O quarto axioma postula que a inclusão de uma terceira opção não altera a predileção por uma das opções previamente definidas. Sendo assim, se M é preferível a L , ($M > L$), para qualquer N :

$$\alpha L + (1 - \alpha)N < \alpha M + (1 - \alpha)N \quad (8)$$

Dados os axiomas apresentados, tem-se o seguinte teorema:

M é preferido a L , $M > L$, se:

$$E(u(L)) < E(u(M)) \quad (9)$$

Onde $E(u)$ corresponde as probabilidades multiplicadas pelo resultado de cada evento. Sendo assim, se a expectativa da utilidade da opção M for superior à expectativa de utilidade da opção L, tem-se como escolha racional a opção M.

Convém lembrar que o modelo é normativo, sugerindo como o ser humano deveria agir para ser considerado racional e maximizador da sua riqueza. Ou seja, foi postulada a ideia de como agir, mas não se mencionou nem testou qual o comportamento dos indivíduos diante de situações envolvendo escolhas. Von Neumann e Morgenstern (1953, p.30) também lembraram que nem todos os sujeitos estão completamente informados nem são capazes de realizar todas as operações estatísticas e matemáticas necessárias. Porém, no estudo em questão, optou-se por não abordar o tema, cabendo tal atividade a estudos futuros por parte de outros autores. No entanto, foi comum ao longo dos anos a pressuposição de racionalidade através deste modelo sem a devida verificação de sua aplicabilidade.

Conforme explicado no modelo, considera-se como racional a escolha maximizadora do resultado da multiplicação da probabilidade de ocorrência do evento pelo resultado financeiro de tal ocorrência.

No entanto, Allais (1953), ao realizar um teste empírico envolvendo diferentes apostas, indicou a existência de inconsistências empíricas nos axiomas propostos na teoria de Von Neumann e Morgenstern (1953) acerca da racionalidade dos indivíduos. Caso análogo ocorreu no estudo de Ellsberg (1961), designado como paradoxo de Ellsberg. Nos dois casos, os estudos indicam a violação dos axiomas do modelo de Von Neumann e Morgenstern (1953).

A implicação das evidências de que os indivíduos não seguem o modelo de racionalidade maximizadora da utilidade está no fato de que estes servem como premissas tanto do CAPM quanto da HME.

Ao evidenciar a existência de comportamento distinto do proposto pelos modelos vigentes, surgem duas opções; incluem-se nos modelos existentes novos fatores capazes de captar os vieses comportamentais ou substituem-se por completo os referidos modelos.

Seja qual for o rumo da pesquisa no campo das finanças, tem-se a necessidade de compreender quais são os fatores que condicionam a tomada de decisão do investidor, sob pena de se propor modelos de precificação ineficazes.

2.3 Teoria da perspectiva

Em substituição à pressuposição de racionalidade utilitária do investidor, Kahneman e Tversky (1979) propuseram a *Prospect Theory*, que consiste na apresentação de um modelo que contesta a teoria da utilidade esperada. Trata-se de um modelo descritivo baseado nas características observadas através das decisões dos indivíduos que substitui a função de probabilidade por uma função de ponderação e a função de utilidade é substituída por uma função de valor.

O objetivo foi apresentar uma nova teoria, *Prospect Theory*, juntamente com uma crítica à teoria da utilidade esperada enquanto modelo de tomada de decisão. O trabalho de Kahneman e Tversky (1979, p. 263, tradução nossa)

[...] consistiu em descrever várias classes de problemas de escolha nos quais as preferências violam sistematicamente os axiomas da teoria da utilidade esperada. À luz destas observações nós argumentamos que a teoria da utilidade, como é comumente interpretada e aplicada, não é um modelo descritivo adequado e propomos uma alternativa que dá conta da escolha sob risco.

Uma das características metodológicas do estudo foi o uso de questões hipotéticas como forma de se obter possíveis escolhas por parte dos indivíduos que compuseram a amostra, representada por estudantes universitários de três países. O fato de não se saber ao certo qual seria a decisão em uma situação real ou em um experimento em laboratório é uma desvantagem do estudo inclusive mencionada pelos autores. Para contornar o fato, foi considerada a premissa de que as pessoas saberiam como se comportar em uma situação real e responderiam de acordo, além da simplicidade do referido método. Também há menção à imperfeição de outros métodos adotados como forma de justificar o tipo de pesquisa utilizado.

Segundo os autores, as pessoas subvaloram resultados que apresentam probabilidades inferiores a um em comparação com eventos certos - efeito certeza. Este efeito aumenta a aversão ao risco em situações que envolvem ganhos e aumenta a propensão ao risco em situações de perda.

Também foi observado que as pessoas descartam informações que são comuns a todas as perspectivas disponíveis - efeito isolamento -, que leva a preferências inconsistentes quando a mesma opção é oferecida de diferentes formas.

Além do efeito certeza e do efeito isolamento, foi percebido que os indivíduos são avessos ao risco em condições de ganho e propensos ao risco em condições de perda, chamado efeito reflexão.

Metodologicamente, foram apresentados dois pares de perguntas envolvendo situações semelhantes em cada par, mas com a diferença de que na primeira delas fora oferecida a possibilidade de um ganho certo. Como resultado, as duas questões, que ofereceriam resultados semelhantes caso os respondentes utilizassem a teoria da utilidade esperada, obtiveram preferências invertidas.

A constatação mais importante a se perceber com o resultado observado é que o axioma da independência foi violado, pois a preferência mudara de acordo com a probabilidade de ocorrência do evento.

Quando as probabilidades de ganho são altas, os indivíduos optam por aquela cuja probabilidade de ganho é maior. Quando as probabilidades são muito baixas, a maioria das pessoas prefere o resultado que oferece o maior ganho (KAHNEMAN e TVERSKY, 1979, p. 267). Os autores consideram qualquer resultado com probabilidade superior a zero como possível e os eventos com probabilidade elevada como prováveis. Desta forma, os indivíduos se comportam de maneira distinta diante de eventos prováveis ou possíveis.

Através da análise das respostas, percebe-se uma inversão nas preferências de acordo com a perspectiva, evidenciando o efeito reflexão. Tal ocorrência viola a condição de racionalidade proposta por Von Neumann e Morgenstern (1953). Na referida pesquisa, os respondentes passaram de uma postura de aversão ao risco em situações envolvendo ganhos para uma postura de propensão ao risco diante de perdas.

Também foi verificado o efeito isolamento. Nas questões apresentadas de maneira distinta, mesmo com probabilidades iguais, os estudantes optaram por alternativas opostas. Se os indivíduos seguissem o padrão de racionalidade proposto por Von Neumann e Morgenstern (1953), o padrão de resposta deveria ser semelhante ao observado anteriormente. Porém, os indivíduos isolaram as duas situações e a maioria optou pela alternativa com ganho certo na segunda fase.

Após apresentados os problemas envolvendo os padrões de comportamento observados nos indivíduos e o que foi postulado pela teoria da racionalidade utilitária, foi apresentada a teoria da perspectiva, que divide a tomada de decisão diante de opções em duas fases; edição e valoração.

A teoria prospectiva distingue duas fases no processo de escolha: uma fase inicial de edição e uma fase subsequente de avaliação. A fase de edição consiste em uma análise preliminar das perspectivas oferecidas, o que muitas vezes gera uma representação simples dessas perspectivas. Na segunda fase, as perspectivas editadas são avaliadas e a perspectiva de maior valor é escolhida. (KAHNEMAN e TVERSKY, 1979, p.274, tradução nossa)

As principais operações da fase de edição são a codificação, a combinação, a segregação, o cancelamento, a simplificação e a detecção de dominância.

Na codificação, as pessoas normalmente percebem os resultados como ganhos ou perdas e não como os estados finais de riqueza. Na operação da combinação, as perspectivas podem ser combinadas com o objetivo de simplificação conforme apresentado no efeito isolamento nos resultados da pesquisa. Na segregação, duas alternativas que compartilham de determinado valor em comum podem ser decompostas em uma alternativa de ganho certo e o risco extra (KAHNEMAN e TVERSKY, 1979, p.274).

As operações cancelamento, simplificação e detecção de dominância são aplicadas a pelo menos duas perspectivas de uma vez.

O cancelamento apresenta situação análoga ao efeito isolamento, mas no contexto de mais de uma perspectiva. Neste caso são oferecidos dois conjuntos de possibilidades para que o indivíduo opte por um e fique exposto ao efeito da aleatoriedade. Consiste na desconsideração do que é comum a todas as alternativas, seja uma fase comum ou alternativas iguais dentre os conjuntos apresentados. A simplificação consiste no arredondamento das probabilidades ou dos resultados. Também é fruto da simplificação a desconsideração de resultados muito improváveis e a detecção de dominância consiste em procurar por alternativas que não maximizem o resultado final (KAHNEMAN e TVERSKY, 1979, p.274).

Tais operações de edição têm como propósito a facilitação do processo decisório, sendo realizadas sempre que possível, mas que levam a vieses que não necessariamente maximizam o resultado.

Na etapa de valoração, espera-se a maximização do valor por parte do interessado. Na teoria proposta, à cada probabilidade p é adicionado um peso π e é adicionado um peso v aos resultados de cada evento possível. A equação resultante é apresentada a seguir:

$$V(x, p; y, q) = \pi(p)v(x) + \pi(q)v(y) \quad (10)$$

Nota-se a inclusão dos pesos, π e v , em comparação com a função de utilidade anteriormente apresentada.

A equação acima apresentada trata de situações nas quais há possibilidade de ganho ou perda. Nas situações em que se verificam resultados positivos/negativos para todas as alternativas, tem-se a seguinte equação:

$$V(x, p; y, q) = v(y) + \pi(p)[v(x) - v(y)] \quad (11)$$

Verifica-se acima o efeito segregação no qual duas alternativas de ganho ou perda certa são combinadas em um evento certo somado às diferenças entre os demais resultados.

2.3.1 A função de valor

Para os autores, o indivíduo leva em conta as diferenças na riqueza mais do que o resultado final dos seus ativos, contrariando Bernoulli (1954). No entanto, ressalta-se que a função de valor da teoria da perspectiva depende de um referencial inicial. Este ponto é crítico para a compreensão das teorias porque não se propôs em nenhuma das teorias que se observasse apenas o valor do ganho ou da perda estritamente, mas sim que se apresentasse um padrão de riqueza inicial (BERNOULLI 1954) ou um referencial de valor (KAHNEMAN e TVERSKY, 1979). Para as duas teorias, pode-se entender que uma mudança de R\$ 200,00 é compreendida de maneira diferente caso o ponto de referência seja R\$ 100,00 ou R\$ 10.000,00.

A função de valor é normalmente côncava para mudanças acima do ponto de referência:

$$v''(x) < 0, \text{ para } x > 0 \quad (12)$$

E convexa para mudanças abaixo do ponto de referência:

$$v''(x) > 0, \text{ para } x < 0 \quad (13)$$

Dados os formatos das curvas, conclui-se que perdas têm efeito mais pronunciado do que ganhos. A representação gráfica da função de utilidade deixa de ser uma reta e passa a ter o comportamento observado na figura abaixo:

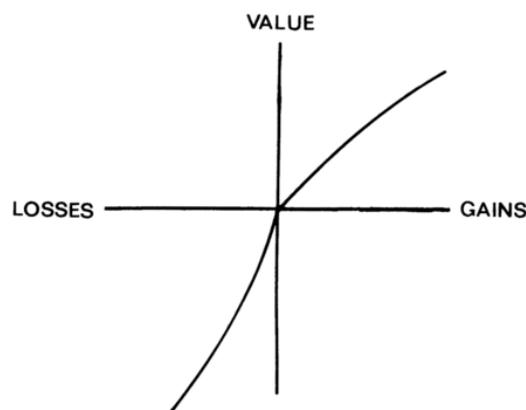


Figura 2 (2): Função de valor proposta pela teoria da perspectiva.

Fonte: Kahneman e Tversky (1979, p. 279)

Kahneman e Tversky (1979, p. 279, tradução nossa) concluem afirmando:

Em resumo, propusemos que a função valor é (i) definida sobre desvios em relação ao ponto de referência; (ii) geralmente côncava para ganhos e comumente convexa para perdas; (iii) mais íngreme para perdas do que para ganhos.

Acerca das funções de valor, observou-se que os indivíduos se sujeitam a maiores riscos com o objetivo de reduzir perdas. Em situações nas quais o objetivo é a obtenção de um determinado ganho, o indivíduo tende a ser mais conservador do que naquelas em que há certeza em relação a uma perda.

2.3.2 A função de ponderação

A teoria da perspectiva postula que o indivíduo não observa apenas os resultados no cálculo, mas também utiliza uma função que pondera os resultados (π). Para pequenas probabilidades, cujos valores são inferiores a 0,001, tais probabilidades são superestimadas, indicando um $\pi(p) > p$. Tal característica leva a uma curva côncava nas situações que representam ganhos e convexa nas situações envolvendo perdas.

Outra característica da função π é a soma das probabilidades ponderadas pela função $\pi(p)$, que é normalmente inferior a um, conforme apresentado na equação abaixo:

$$\pi(p) + \pi(1-p) < 1 \quad (14)$$

Kahneman e Tversky (1979) chamaram o efeito de subcerteza. Este efeito faz com que haja predileção pelos eventos certos.

Dado que há superestimação das probabilidades dos eventos com pequenas probabilidades e dado que a soma das probabilidades ponderadas é inferior a um, tem-se uma relação inversa entre os valores de π e p . Tal característica foi chamada subproporcionalidade e pode ser melhor compreendida na figura abaixo:

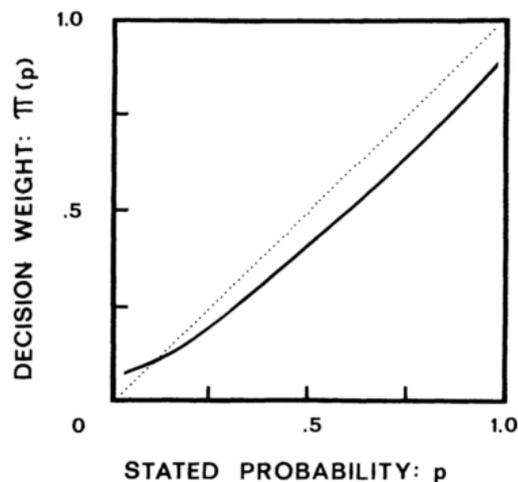


Figura 3 (2): Função de ponderação
 Fonte: Kahneman e Tversky (1979, p. 283)

Acerca da função de probabilidade, os autores concluem que

Porque as pessoas são limitadas em sua capacidade de compreender e avaliar as probabilidades extremas, eventos altamente improváveis são ignorados ou superdimensionados e a diferença entre alta probabilidade e certeza é negligenciada ou exagerada. Consequentemente, π não se comporta de maneira adequada perto dos limites (KAHNEMAN e TVERSKY, 1979, p. 283).

Como avanço em relação às demais teorias sobre racionalidade do comportamento humano, tem-se um modelo descritivo para o tema. Os modelos anteriores foram de caráter normativo, indicando como o indivíduo deveria agir e não apresentaram evidências empíricas quanto ao grau de assertividade em situações práticas. A teoria da perspectiva, ao apresentar evidências do comportamento humano mais do que suposições de como este deveria ser, oferece maior chance de acerto na busca pelo modelo mais adequado de precificação de ativos.

2.3.3 Vieses comportamentais

Há estudos sobre as implicações da *Prospect Theory* nas finanças, refletidas na busca por vieses comportamentais que façam com que os indivíduos abandonem o padrão de racionalidade proposto no modelo clássico de finanças. O primeiro deles, o efeito dotação, proposto por Thaler (1980), estabelece que o investidor tende a superestimar o valor dos ativos que possui, sendo o efeito maior nos bens para usufruto próprio e naqueles cujos valores não são facilmente verificáveis. A implicação aqui observada é que perder um bem que se possui é mais significativo do que ganhar algo ainda não possuído, conforme

verificado no efeito reflexão. Tal efeito foi verificado por Hwang e Rubesam (2013), que identificaram uma assimetria na resposta a informações públicas por parte dos investidores.

O segundo efeito, conhecido como efeito disposição, proposto por (SHEFRIN; STATMAN, 1984), propõe que o indivíduo tende a ficar mais tempo com ações que apresentam perdas enquanto este tende a vender as ações que estão com algum tipo de ganho. Li e Yang (2013) propuseram um modelo para precificar tal viés. Grinblatt e Han (2005) propuseram uma variável que serviria de *proxy* para identificar o referido efeito.

O terceiro efeito envolve o excessivo prêmio de risco do mercado de ações quando comparado aos mercados de títulos. Benartzi e Thaler (1995) sugeriram que os investidores costumam projetar os seus investimentos em períodos de tempo muito curtos, que, aliados à aversão a perdas, levam à sensibilidade em períodos de perdas do mercado.

Lobão (2012, p 3045) sintetiza outros enviesamentos que foram observados ao longo dos anos, como o da contabilidade mental, que leva os investidores a agrupar seus ativos em categorias não relacionadas às características destes, o excesso de confiança – *overconfidence*, também observado por Chuang e Lee (2006), que faz com que os investidores creiam que suas aptidões cognitivas sejam superiores, o conservadorismo, que leva o indivíduo a dar demasiada importância a informações passadas em detrimento de informações atuais, o enviesamento de disponibilidade, que leva o indivíduo a atribuir importância às informações mais fáceis de recordar, de retrospectiva, que leva o indivíduo a crer que os acontecimentos passados foram mais previsíveis do que realmente foram, a auto atribuição leva o indivíduo a atribuir o seu sucesso a fatores próprios e o fracasso a fatores externos e a ilusão de controle, que leva o indivíduo a crer que possui controle sobre fatores que não pode influenciar.

Chuang e Lee (2006) verificaram a hipótese de *overconfidence* através de quatro implicações testáveis. A primeira delas foi que, uma vez que se encontram excessivamente confiantes, os investidores reagem de maneira excessiva a informações privadas e de maneira insuficiente a informações públicas (DANIEL; HIRSHLEIFER; SUBRAHMANYAM, 1998). A segunda implicação testável estaria na ideia de que ganhos de mercado levariam, uma vez excessivamente confiantes, os investidores a operar de maneira agressiva nos períodos seguintes (GERVAIS; ODEAN, 2001). A terceira envolve a excessiva volatilidade dos mercados em caso de excesso de confiança (GERVAIS; ODEAN, 2001). Por fim, investidores confiantes investiriam em ativos de maior risco (ALPERT e RAIFFA, 1982).

Foram encontradas evidências que dariam suporte às quatro suposições propostas pelos autores.

Schmeling (2009) avaliou o sentimento dos investidores de dezoito países industrializados. As evidências indicaram que sentimentos negativos são preditores do desempenho dos mercados entre os países e que o fator sentimento no retorno das ações ocorre com maior intensidade em países com instituições mais frágeis e em mercados com comportamento de manada.

Baker e Wurgler (2006) combinaram, através de uma análise fatorial, as variáveis disponíveis na literatura existente de finanças comportamentais e criaram uma variável latente chamada de sentimento do investidor.

A hipótese do modelo consistia no fato de que as ações de precificação mais subjetiva, como as de empresas com prejuízo contábil, não distribuidoras de dividendos, com poucos anos de existência, seriam suscetíveis a erros de avaliação em decorrência de vies comportamental. Os resultados indicaram que, em mercados otimistas, as ações de difícil valoração subiam mais que o seu valor fundamental, levando os compradores a perdas futuras. Em esperada oposição, as ações de difícil valoração costumavam ficar subavaliadas em tempos de pessimismo no mercado, gerando ganhos superiores à média do mercado dentro daquele patamar de risco.

Considerando os desvios em relação ao pressuposto de racionalidade utilitária, considera-se, neste trabalho, a ideia de incluir uma variável capaz de identificar os desvios apresentados, chamada ISBR, índice de sentimento do mercado de capitais brasileiro.

2.4 Índices de sentimento

Nesta seção, são apresentados os fundamentos e os estudos envolvendo os índices de sentimento, bem como as definições que explicam o termo e suas características. Inicialmente, são apresentados os fundamentos da avaliação de ativos e seus vieses, passando pelos tipos de indicadores e finalizando com a apresentação de estudos acerca do tema. O conceito de sentimento envolve a propensão à geração de vieses na precificação dos ativos por parte dos investidores em virtude de otimismo ou pessimismo na projeção do retorno futuro dos ativos.

Para compreender a razão da utilização de índices de sentimento, deve-se começar observando que a precificação de determinado ativo se dá pelo valor presente da expectativa

dos seus fluxos de caixa futuros, cuja representação se dá através da equação (MILLER; MODIGLIANI, 1961) (FAMA; FRENCH, 2015):

$$V_0 = \sum_{t=0}^{\infty} E(X_t - I_t) / (1+p)^{t+1} \quad (15)$$

Onde V representa o valor do ativo, X representa o lucro, I representa a necessidade de investimento, também representada pela variação no patrimônio líquido da entidade entre os períodos t e $t-1$, e p representa o custo do capital.

Segundo a teoria clássica de finanças, o investidor é racional e há consenso quanto ao valor de cada ativo, devendo, com base nas informações disponíveis, projetar o retorno e a necessidade de investimento para cada um dos períodos futuros e, em seguida, trazer o resultado ao valor presente, chegando assim ao valor justo do referido ativo. Percebe-se aqui o caráter subjetivo de tal expectativa, que pode levar a vieses de precificação caso o investidor não apresente o padrão de racionalidade sugerido.

Tomando por base a ideia e as evidências de que o investidor não segue sempre o padrão de racionalidade proposto por Von Neumann e Morgenstern (1953) e que age de maneira distinta em situações nas quais há possibilidades de ganho ou perda (KAHNEMAN; TVERSKY, 1979), espera-se que haja diferentes padrões de comportamento que violem os pressupostos da teoria clássica na precificação de ativos. Tal condição leva o mercado a apresentar certa previsibilidade em determinados momentos, o que poderia ser usado para aumentar retornos sem o correspondente aumento dos riscos.

A evidência de que o comportamento muda diante da forma como é apresentado o contexto, ao ser combinada com a ideia de que sentimentos como otimismo ou pessimismo podem ser interpretados como criadores de tal contexto, levam o indivíduo e o mercado a avaliar os ativos de maneira equivocada. Baker e Wurgler (2006) definem o termo sentimento do investidor através de duas possibilidades distintas, sendo a primeira definida como a propensão a especular e a segunda como o otimismo ou pessimismo envolvendo o mercado de ações. Nas duas situações, há certa dificuldade em arbitrar, seja por limitações legais, seja pela subjetividade em avaliar o preço justo das ações em virtude de características próprias como ausência de lucros, histórico recente ou incertezas no mercado.

A fim de identificar o sentimento, duas abordagens são adotadas. A abordagem conhecida como *bottom-up* foca no comportamento do investidor em si e como este reage em situações específicas. Já a abordagem foco deste estudo, chamada *top-down*, observa o sentimento através do valor agregado do mercado e seu efeito no retorno dos ativos.

Como referência da abordagem *bottom-up*, cita-se o trabalho de Barberis, Shleifer e Vishny (1998), que criaram um modelo de comportamento de um investidor com padrão de raciocínio que leva a desvios a partir de vieses como conservadorismo ou representatividade. Os autores apresentam o modelo no trecho a seguir:

No nosso modelo, os resultados de um ativo seguem o padrão *random walk*. No entanto, o investidor não sabe disso. Na verdade, ele acredita que o comportamento do resultado de uma firma se move entre dois 'estados' ou 'regimes'. No primeiro estado, os resultados são revertidos à média. No segundo estado, segue uma tendência, significando que seguem no mesmo sentido do momento anterior. (BARBERIS; SHLEIFER; VISHNY, 1998, p. 309, tradução nossa)

O modelo consiste em projetar o provável comportamento do investidor no processo de *valuation* de um ativo ao determinar se esse acredita que o resultado futuro seguirá um padrão de reversão à média ou se seguirá no mesmo sentido da variação ocorrida na última divulgação. Para tanto, toma por base a fórmula de Bayes, verificando a probabilidade de se estar no estado um ou dois a partir da sequência de resultados anteriores.

Conforme apresentado e visto na descrição do estudo supracitado, na estratégia *bottom-up*, mensura-se o sentimento a partir da análise do comportamento provável de um indivíduo, chegando ao desvio no mercado.

De forma inversa, na abordagem *top-down*, observa-se o comportamento do mercado para extrair informações sobre o comportamento dos indivíduos que o compõem. Sendo assim, nos momentos de euforia, esperam-se avaliações com viés otimista e consequente sobrepreço dos ativos, nos momentos de pânico, esperam-se avaliações com viés pessimista e consequente subavaliação dos referidos ativos. Como consequências, tem-se um retorno acima da média para os ativos precificados durante um período de pessimismo generalizado e um retorno abaixo da média a partir de períodos de otimismo nos mercados.

Sugerindo tal possibilidade, Baker e Wurgler (2006) estudaram como o sentimento do investidor afetou o retorno dos ativos no mercado de capitais norte-americano entre 1961 e 2002 a partir do agregado de variáveis, chegando a conclusão de que o índice utilizado se comportou conforme esperado principalmente para ações de empresas pequenas, com alta volatilidade, oriundas de empresas jovens, não lucrativas, não pagadoras de dividendos e na iminência de falência.

A utilização de uma variável latente como *proxy* para o sentimento do investidor se deu pela dificuldade de observação do sentimento diretamente. Para compor tal variável, os autores utilizaram um grupo de *proxies* disponíveis, conforme afirmaram:

Uma vez que não há nenhuma *proxy* perfeita ou não controversa para o sentimento do investidor, nós consideramos um número de *proxies* utilizadas recentemente para formar um índice composto de sentimento baseado no primeiro componente principal [de uma análise fatorial]. Para diminuir a verossimilhança entre as *proxies* e o risco sistemático, nós também formamos um índice baseado no sentimento que fora ortogonalizado com algumas condições macroeconômicas (BAKER; WURGLER, 2006, p. 1646, tradução nossa).

Seguindo a ideia de que o sentimento do investidor não é diretamente observável, tem-se a necessidade de evidenciação do componente em comum das *proxies* utilizadas. Para tanto, estabeleceu-se como primeira hipótese desta tese: Há um componente principal obtido a partir das *proxies* de sentimento observadas. A confirmação ou o falseamento desta hipótese é importante na tentativa de responder à pergunta de pesquisa proposta.

Para compor o índice de sentimento de Baker e Wurgler (2006), foram escolhidas seis variáveis. A primeira foi o desconto no valor das cotas dos *close end funds*, que é obtido pela diferença entre a soma das cotas do fundo e o valor das ações que o compõem. A ideia reside no fato de que, dado que tais fundos não vendem as ações que fazem parte da sua carteira, um investidor pessimista se desfaria de tais cotas pagando um prêmio por acreditar que, no futuro, o valor dos ativos iria cair.

A segunda variável escolhida foi o *turnover* das ações da NYSE, que é calculado a partir da divisão entre o volume transacionado e o número médio de ações em bolsa. A ideia é que os investidores otimistas tendem a operar mais vezes.

O número de IPOs e o retorno médio do primeiro dia de negociações após as IPOs correspondem a terceira e quarta variáveis por serem consideradas reflexos de entusiasmo do mercado. Considera-se que um número elevado de IPOs sugere otimismo enquanto um número baixo reflete apreensão.

A relação entre aumento de capital próprio e aumento total de capital (próprio e de terceiros) fora considerada outra forma de captar o sentimento do investidor. Segundo os autores, um período de elevado aumento de capital próprio é seguido por um período de baixos retornos de mercado.

A última *proxy* foi o que os autores chamaram de prêmio de dividendos, que corresponde à diferença entre o *Book-to-Market* de pagadoras e não pagadoras de dividendos.

Os autores então realizaram uma análise fatorial das variáveis escolhidas, sendo o primeiro componente utilizado como *proxy* de sentimento para a análise, resultando no seguinte índice (BAKER; WURGLER, 2006, p. 1657):

$$\begin{aligned} SENTIMENT_T = & -0,241 CEFD_t + 0,242 TURN_{t-1} + 0,253 NIPO_t + 0,257 RIPO_{t-1} \\ & + 0,112 S_t - 0,283 P_{T-1}^{D-ND} \end{aligned} \quad (16)$$

Onde CEFD corresponde ao desconto no valor das cotas dos *close end funds*, TURN representa o *turnover* dos últimos doze meses, NIPO representa o número de IPOs, RIPO corresponde ao retorno médio no primeiro dia de negociações, S indica o aumento de capital próprio em relação ao aumento total de capital e P representa o logaritmo da diferença entre o *book-to-market* das pagadoras e não pagadoras de dividendos.

Após ortogonalizar o modelo pelo índice de produção industrial, consumo de bens duráveis, bens não duráveis, serviços e uma variável *dummy* para recessões, obteve-se o seguinte índice (BAKER; WURGLER, 2006, p. 1658):

$$\begin{aligned} SENTIMENT_T = & -0,198 CEFD_t + 0,225 TURN_{t-1} + 0,234 NIPO_t + 0,263 RIPO_{t-1} \\ & + 0,211 S_t - 0,243 P_{T-1}^{D-ND} \end{aligned} \quad (17)$$

A variável índice foi incluída no modelo de três fatores proposto por Fama e French (1992) com a respectiva classificação em decis proposta. Os resultados indicaram que as empresas que apresentavam maior complexidade na mensuração de valor sofriam maior influência do índice de sentimento do mercado, notadamente as empresas pequenas, com ações que apresentavam alta volatilidade, jovens, não lucrativas, não pagadoras de dividendos e na iminência de falência. Desta forma, o desenvolvimento de um índice de sentimento permitiu um ajuste na valoração das empresas para corrigir erros de precificação fruto de vieses comportamentais por parte dos investidores.

Conforme mencionado, esta tese teve como primeira hipótese a ideia de que há um componente principal obtido a partir das *proxies* de sentimento observadas. Tal proposta foi adotada no trabalho de Yoshinaga (2009), que criou um índice de sentimento para o mercado brasileiro tomando por base o modelo proposto por Baker e Wurgler (2006). Como diferencial teórico desta tese em comparação com os trabalhos citados, há a adição de uma variável *proxy* que mensure o contágio entre os sentimentos brasileiro e americano.

A ideia de contágio foi proposta por Baker, Wurgler e Yuan (2012) e testada em seis mercados, que consistiu na evidênciação de um componente em comum de todos os índices de sentimento.

Para esta tese, pretendeu-se verificar se o índice de sentimento norte-americano possui variância em comum com as *proxies* do índice brasileiro de sentimento, sugerindo a existência de contágio. Tal proposição é estabelecida na segunda hipótese deste trabalho: O

índice de sentimento norte-americano está correlacionado com as *proxies* de sentimento do mercado brasileiro.

As variáveis adotadas por Yoshinaga (2009) seguiram, em sua maioria, o que fora proposto por Baker e Wurgler (2006), sendo mantidas as variáveis S, NIPO, TURN e DIV. Nota-se a exclusão das variáveis CEFD e RIPO, referentes ao desconto dos *close end funds* e ao retorno no primeiro dia de negociação respectivamente. A exclusão de variáveis se deu em virtude da baixa liquidez dos *close end funds* no Brasil e dos períodos sem IPOs.

O estudo incluiu uma nova variável, *trading index*, utilizada na análise técnica para identificar se há maior proporção de altas ou de baixas no mercado. O indicador, conhecido como TRIN, é utilizado para identificar se o mercado está com maior força compradora, *bullish*, ou com maior força vendedora, *bearish*. A equação de cálculo do indicador é apresentada abaixo

$$TRIN = \frac{\frac{\text{Número de ações com preço em alta}}{\text{Número de ações com preço em queda}}}{\frac{\text{Número de ações com volume em alta}}{\text{Número de ações com volume em queda}}} \quad (18)$$

Caso o indicador apresente valor superior a um, o mercado está pessimista ou com maior força vendedora. Caso apresente valor inferior a 1, o mercado está otimista ou com maior força compradora.

O indicador de sentimento foi então calculado e representado na equação:

$$SENT = 0,3941 S_t + 0,5574 NIPO_t + 0,4796 TURN_{T-1} - 0,4802 P_{T-1} - 0,2708 TRIN_t \quad (19)$$

Os autores também ortogonalizaram as variáveis para expurgar o efeito dos ciclos de negócios, que poderiam atuar como variáveis espúrias e comprometer a validade do estudo. Foram incluídas na ortogonalização o PIB - Produto Interno Bruto - do país, da agricultura, da indústria, dos serviços, uma *dummy* para variações positivas ou negativas no PIB do país e uma variável *dummy* para variações positivas ou negativas na taxa selic.

Com a ortogonalização das variáveis, o indicador apresentou a seguinte configuração:

$$SENT^\perp = 0,4154 S_t + 0,5239 NIPO_t + 0,4635 TURN_{T-1} - 0,5030 P_{T-1} - 0,2918 TRIN_t \quad (20)$$

Após calcular o índice de sentimento, a amostra foi separada em decis de acordo com as características valor de mercado, risco e tempo de listagem em bolsa. Após a separação em decis, foram calculados os retornos médios após períodos de sentimento positivo e negativo.

A conclusão indicou que, após períodos de otimismo, o retorno das ações cuja precificação exigiu mais subjetividade foi inferior ao retorno observado após períodos de pessimismo. Este resultado segue em linha com o observado por Baker e Wurgler (2006).

Lemmon e Portniaguina (2006) e Schmeling (2009), adotaram o índice de confiança do consumidor como *proxy* para o sentimento do investidor. Este índice se mostra disponível em vários países, possui uma série histórica longa e serve de guia para a tomada de decisão de vários agentes.

As evidências mencionadas sugerem que o sentimento altera o apreçamento dos ativos do mercado, sendo possível auferir ganhos extraordinários em virtude da inclusão de uma variável que mensure tal fenômeno.

A possibilidade de ganhos extraordinários leva à terceira hipótese deste estudo: O índice de sentimento proposto melhora o poder de explicação dos retornos das ações negociadas no mercado de capitais brasileiro. Ao afirmar que o sentimento melhora o poder de explicação, tem-se a proposição de um novo fator nos modelos de apreçamento, sendo possível melhorar a eficiência e o poder de explicação dos referidos modelos.

Além da alteração nos retornos dos ativos em virtude do sentimento do investidor, Baker, Wurgler e Yuan (2012) encontraram evidências de contágio do sentimento de diferentes países. Possíveis evidências de contágio levaram à quarta hipótese desta tese: A inclusão da variável representativa do contágio, representada pelo índice de sentimento norte-americano, melhora o poder de explicação do índice de sentimento do investidor brasileiro.

A combinação das hipóteses propostas levou à proposição principal desta tese, que envolveu a criação de um índice de sentimento que considere em sua composição as evidências de contágio entre sentimentos.

Tem-se, como proposta resposta à pergunta de pesquisa sugerida neste estudo, a ideia de elaboração de um índice de sentimento do investidor a partir dos modelos de sentimento apresentados, composto por variáveis disponíveis ao investidor do mercado de capitais brasileiro. Como inovação em comparação com Baker e Wurgler (2006) e Yoshinaga (2009), tem-se a pressuposição de contágio entre sentimentos de diferentes países, representada pelo índice de sentimento norte-americano. O tratamento dado às variáveis está descrito na terceira seção desta tese, procedimentos metodológicos, a seguir.

3 Procedimentos metodológicos

Esta tese teve por objetivo verificar o efeito no retorno das ações da BM&FBOVESPA quando considerado o sentimento do mercado de capitais norte-americano na propensão à geração de vieses dos investidores brasileiros no período de 2006 a 2015. Para atingir o objetivo proposto, foi realizado um trabalho quantitativo com fonte de dados secundária em duas etapas. A primeira etapa consistiu na elaboração de dois índices de sentimento a partir de uma análise de componentes principais – ACP – conforme descrito na subseção 3.3.1. A segunda etapa consistiu no teste dos índices obtidos, com os procedimentos descritos na subseção 3.4, através da utilização de dados em painel tomando por base o método adotado por Fama e French (2015) no modelo de cinco fatores.

Esta seção está dividida em cinco subseções, quais sejam: (1) Descrição das variáveis utilizadas no desenvolvimento do estudo; (2) Descrição da coleta e organização dos dados; (3) Apresentação do modelo utilizado na literatura acadêmica envolvendo a formação de variáveis latentes ou fatores; (4) Descrição do modelo econométrico utilizado nos testes das variáveis obtidas; (5) Apresentação das hipóteses propostas.

3.1 Delineamento, população, amostra e variáveis

Este trabalho foi desenvolvido tomando por base o paradigma pós-positivista e suas consequentes ramificações de natureza ontológica, epistemológica e metodológica. A ontologia adotada tem por foco o indivíduo, seguindo a tradição observada nas áreas das finanças clássicas das finanças comportamentais.

Quanto à epistemologia, foi feito um estudo de natureza exploratória, em consonância com os trabalhos publicados nas finanças comportamentais envolvendo a temática sentimento do investidor.

Quanto aos métodos, trata-se de levantamento com fonte de dados secundária em duas etapas. A primeira etapa consistiu na elaboração de dois índices de sentimento. O intuito da elaboração de dois índices foi testar as hipóteses III e IV propostas. A segunda etapa consistiu no teste dos índices obtidos através da utilização do teste GRS e de regressões com dados em painel.

A população desta tese compreendeu as ações das empresas de capital aberto no Brasil entre os anos de 2006 e 2015. A amostra coletada envolveu as ações com negociações em bolsa em pelo menos um mês no período mencionado, totalizando trezentas e cinquenta e seis ações pertencentes a duzentas e setenta e cinco empresas.

A apresentação das variáveis é iniciada pela apresentação das *proxies* que compuseram os índices de sentimento, seguidas pelas variáveis utilizadas nos modelos de apreçamento. Por fim, no tópico 3.2, são detalhados os procedimentos de coleta e organização dos dados.

3.1.1 Variáveis do índice de sentimento

A escolha das variáveis do índice de sentimento proposto tomou por base os estudos de Yoshinaga (2009) e Baker, Wurgler e Yuan (2012). O primeiro estudo apresentou uma proposta de índice para o mercado brasileiro, sendo uma adaptação para o mercado nacional do índice proposto por Baker e Wurgler (2006). Ainda sobre o trabalho de Yoshinaga (2009), tem-se também como contribuição para esta tese a seleção das variáveis disponíveis no mercado brasileiro, sendo excluídas aquelas sem liquidez ou enviesadas. O segundo estudo, por sua vez, apresentou a ideia de contágio, que faz parte da proposta desta tese e pode contribuir com a melhoria do poder de explicação dos modelos de precificação de ativos existentes.

Para a obtenção do índice formado a partir dos trabalhos mencionados, dado que o sentimento do investidor não é diretamente observável, fez-se necessário recorrer a um conjunto de *proxies* que o capturaram de maneira indireta. Tais variáveis foram compiladas no quadro abaixo e descritas em seguida.

A partir de tais variáveis é possível extrair uma variável latente para o índice proposto, sendo estas variáveis utilizadas como indicadores do sentimento do mercado brasileiro. Espera-se que estes índices apresentem maior precisão e poder de explicação em comparação com cada *proxy* isolada.

A seguir são detalhadas as características e fundamentadas as escolhas das variáveis utilizadas:

- **Número de IPOs:**

O histórico da quantidade de ofertas públicas iniciais sugere alguma relação com o sentimento do mercado. Em períodos de exuberância costumam haver muitas ofertas, enquanto essas cessam ou escasseiam em períodos de crise. A ideia aqui postulada é de que o

mesmo sentimento que afeta o número de lançamentos afeta também a propensão do investidor a avaliar os ativos de maneira viesada, sendo esperada uma relação positiva entre o sentimento do investidor e o número de IPOs.

$$NIPO_t = \text{Número de IPOs no último semestre} \quad (21)$$

A periodicidade da variável NIPO não é consenso entre os trabalhos pesquisados. Também não foi observada justificativa explícita por parte dos autores que fizeram uso de diferentes intervalos na variável em questão. Enquanto Baker e Wrugler (2006) utilizaram o número de ofertas dos últimos doze meses, Yoshinaga (2009) utilizou o número de IPOs dos últimos três meses. Nesta tese, optou-se por incluir a média móvel aritmética do número de IPOs do último semestre, representada através da equação:

$$MM_t = \frac{\sum_{i=0}^5 p_{t-i}}{6} \quad (22)$$

Onde p_t representa o número de ofertas no mês t .

Os dados da quantidade de ofertas públicas são disponibilizados diretamente pela Bolsa de Valores de São Paulo - Bovespa.

- **Retorno dos IPOs:**

O retorno obtido no primeiro dia de negociações pós-IPO também sugere relação com vieses na avaliação dos ativos, sendo a relação positiva entre o retorno e o sentimento. Grandes retornos iniciais sugerem erros de precificação por parte do otimismo dos investidores inexperientes ou dos especuladores. Analogamente, resultados negativos indicam sentimentos que levam à subprecificação dos ativos. O cálculo do indicador se deu através da equação:

$$RIPO_t = \frac{\sum (P_{fi} - P_{oi})}{n} \quad (23)$$

Onde P_{fi} representa o preço de fechamento no primeiro dia de negociações pós-IPO, P_{oi} representa o preço definido na IPO e n representa o número de IPOs no período.

Seguindo o padrão adotado na variável NIPO, foram considerados os retornos das ofertas públicas dos últimos seis meses no cálculo dos valores que compuseram a base de dados da variável. As informações sobre os preços estabelecidos nas ofertas e os valores de

fechamento do primeiro dia de negociações são disponibilizadas pela Bovespa em seu *website*.

- **Prêmio de dividendos:**

O prêmio de dividendos é calculado pela relação entre o *book-to-market* das pagadoras e das não pagadoras de dividendos. Pressupõe-se que a busca por pagadoras de dividendos, ou aversão a não pagadoras, se dá em períodos de sentimento negativo por parte dos investidores, uma vez que estes valoram as não pagadoras de maneira excessivamente pessimista. Sendo assim, espera-se um sinal positivo na relação entre sentimento e prêmio de dividendos. A equação é representada por:

$$\text{Div}_{p-np} = \ln BTM_p - \ln BTM_{np} \quad (24)$$

Onde BTM_p corresponde à média ponderada do índice *book-to-market* das empresas pagadoras de dividendos e BTM_{np} das não pagadoras de dividendos.

Os dados referentes ao *book-to-market*, bem como a listagem das empresas pagadoras e não pagadoras de dividendos, foram obtidas na base de dados econômica.

- **Sentimento do mercado norte-americano:**

A variável de sentimento elaborada por Baker e Wurgler (2006) para o mercado norte-americano foi incluída como *proxy* para o contágio de sentimento entre os mercados norte-americano e brasileiro. A suposição de que haja ligação entre os sentimentos externos ao Brasil e que esta ligação leve ao “contágio” no país a partir de fontes externas pode melhorar o poder de explicação do índice de sentimento, indicando um sinal positivo na relação entre o sentimento do mercado americano e do mercado brasileiro. A base de dados está disponível no *website* do professor Jeffrey Wurgler.

- **Índice de confiança do consumidor:**

O índice de confiança do consumidor, medido e disponibilizado pelo IBRE/FGV, procura “gerar informações que reflitam as condições macroeconômicas vigentes e extrair informações de âmbito psicológico não captadas por indicadores econômicos tradicionais, contribuindo desta forma ao aperfeiçoamento dos modelos de previsão econômica” (FGV, 2012, p. 3). A disponibilização dos dados é feita pelo próprio instituto.

Lemmon e Portniaguina (2006) encontraram evidências de que o índice de sentimento do consumidor servira como *proxy* de sentimento do mercado, sugerindo uma relação positiva entre o índice e o sentimento do mercado. Schmeling (2009) utilizou o índice de confiança do

consumidor como *proxy* de sentimento em dezoito países, encontrando evidências análogas aos demais trabalhos, sendo o efeito mais intenso em países que são culturalmente direcionados a comportamentos de manada e sujeitos a *overreaction*.

Quadro 1(3): Descrição das variáveis utilizadas na elaboração dos índices de sentimento

Variável	Operacionalização	Sinal esperado	Referência
Número de IPOs (NIPO)	$NIPO_t = \sum_{i=1}^6 IPO_t$	Positivo	Baker e Wurgler (2006)
Retorno dos IPOs (RIPO)	$RIPO_t = \frac{\sum (P_{fi} - P_{oi})}{P_{oi} \cdot n}$	Positivo	Baker e Wurgler (2006)
Prêmio de dividendos (Div)	$Div_{p-np} = \ln BTM_p - \ln BTM_{np}$	Positivo	Baker e Wurgler (2006)
Índice de confiança do consumidor (ICC)	Série histórica fornecida pela Fundação Getúlio Vargas.	Positivo	Lemmon e Portniaguina (2006), Schmeling (2009)
Sentimento do mercado norte-americano (Contágio)	Série histórica fornecida pelo professor Jeffrey Wurgler.	Positivo	Baker, Wurgler e Yuan (2012)

Fonte: Autoria própria, 2016.

O quadro 2 descreve as variáveis utilizadas e suas respectivas operacionalizações.

3.1.2 Variáveis do modelo de apreçamento

As variáveis do modelo de apreçamento seguiram o modelo de cinco fatores proposto por Fama e French (2015), cujas variáveis dependentes e independentes são apresentadas a seguir.

- **Retorno das carteiras de ações:**

A variável dependente do modelo de cinco fatores, assim como do modelo clássico e do modelo de três fatores, é o retorno das carteiras, calculado através da equação:

$$R_{i,t} = \frac{P_{i,t}}{P_{i,t-1}} \quad (25)$$

Onde P corresponde ao preço médio de fechamento das ações no pregão da carteira i no mês t , já descontada a taxa livre de risco.

Os dados dos preços de fechamento foram obtidos através da base de dados econômica. Foram obtidos os preços de fechamento ajustados pelos dividendos para evitar vieses de variação no preço desses ativos em virtude do pagamento de dividendos.

- **Retorno do mercado:**

A primeira variável independente corresponde ao retorno do mercado, obtida através da subtração do retorno do índice bovespa pela taxa selic do respectivo mês:

$$R_{m,t} = R_{M,t} - R_{f,t} \quad (26)$$

Onde $R_{M,t}$ corresponde ao retorno do *Índice Bovespa no período t* e $R_{f,t}$ corresponde ao valor médio da taxa selic no mês t.

Assim como as demais variáveis do modelo de apreçamento, esta também foi obtida através da base de dados econômica.

- **Tamanho:**

Utilizado para criar o fator SMB, *Small Minus Big*, e para compor as carteiras utilizadas como variáveis independentes, corresponde ao tamanho da empresa. A variável é mensurada através do valor de mercado da empresa. A equação corresponde a:

$$T = \ln VM_{i,t} \quad (27)$$

Onde \ln corresponde ao logaritmo natural e $VM_{i,t}$ valor de mercado do ativo i no período t .

Os dados relacionados aos valores de mercado das empresas foram obtidos na base de dados econômica.

- **Book-to-market:**

Utilizado como segundo fator do modelo, denominado HML, *High Minus Low*, e na elaboração das carteiras utilizadas como variáveis dependentes, o indicador mensura a relação entre o valor contábil do patrimônio líquido e o valor de mercado de determinado ativo. Para composição dos dados desta variável foram coletados os valores na base de dados econômica.

$$B/M = \frac{B_{i,t-1}}{M_{i,t-1}} \quad (28)$$

Onde $B_{i,t}$ corresponde ao valor contábil do patrimônio líquido do ativo i no período $t-1$ e $M_{i,t}$ o valor de mercado do ativo i no período $t-1$.

- **Lucratividade:**

A variável lucratividade é mensurada através do lucro operacional incluindo a despesa financeira dividido pelo patrimônio líquido da empresa, conforme apresentado na equação:

$$L = \frac{LO_{t-1}}{PL_{t-1}} \quad (29)$$

Onde LO corresponde ao lucro operacional e PL ao patrimônio líquido, ambos no ano anterior ao do modelo, ambos obtidos através da base de dados econômica.

- **Investimento:**

A variável investimento corresponde à variação do ativo total entre os dois últimos anos fiscais dividido pelo ativo total do penúltimo ano fiscal, também obtidos através da base de dados econômica, conforme equação:

$$I = \frac{AT_{t-1} - AT_{t-2}}{AT_{t-2}} \quad (30)$$

Onde AT corresponde ao ativo total da empresa.

Quadro 2(3): Descrição das variáveis utilizadas no modelo de teste dos índices de sentimento

Variável	Descrição	Operacionalização	Tipo de variável
Retorno da carteira	Retorno mensal da carteira de ações.	$R_{i,t} = \frac{P_{i,t}}{P_{i,t-1}}$	Dependente
Retorno do mercado	Retorno mensal do mercado.	$R_{m,t} = R_{M,t} - R_{f,t}$	Independente
Tamanho	Tamanho da empresa.	$T = \ln VM_{i,t}$	Independente
<i>Book-to-market</i>	Relação entre valor patrimonial e valor de mercado da empresa.	$B/M = \frac{B_{i,t-1}}{M_{i,t-1}}$	Independente
Lucratividade	Relação entre o resultado operacional da empresa e o seu capital próprio.	$L = \frac{LO_{t-1}}{PL_{t-1}}$	Independente
Investimento	Investimento realizado pela empresa entre o antepenúltimo e o penúltimo ano.	$I = \frac{AT_{t-1} - AT_{t-2}}{AT_{t-2}}$	Independente

Fonte: Autoria própria, 2016.

O quadro 3 acima apresenta e descreve as variáveis utilizadas no modelo de cinco fatores.

3.2 Coleta e organização dos dados

Os dados utilizados na elaboração do índice de sentimento e na montagem do modelo de teste foram obtidos em diferentes fontes. Os dados referentes às ofertas públicas iniciais são disponibilizados pela BM&FBovespa em seu *website*, bem como o histórico de preços das

respectivas ações. O índice *market-to-book*, a lista de pagadoras e não pagadoras de dividendos e as variáveis que compõem o modelo de cinco fatores de Fama e French (2015) foram obtidas na base de dados Economática. O índice de sentimento norte-americano é atualizado e disponibilizado pelo professor Jeffrey Wurgler, PhD, em seu *website*. O histórico do índice de confiança do consumidor foi obtido através do IBRE/FGV, entidade responsável pela publicação do referido índice.

Dado que a série histórica do índice de sentimento do investidor é composta por observações a partir do ano de 2006³ e o ano de 2004 ser referenciado como a volta às IPOs após alguns anos sem ofertas no mercado primário, optou-se por um intervalo de nove anos a partir do mês de julho do ano de 2006 até o mês de junho de 2015.

Este trabalho dividiu-se em duas etapas. A primeira etapa correspondeu à elaboração do índice de sentimento ISBR e o índice de sentimento sem a variável representativa do sentimento norte-americano, sendo utilizado o segundo índice para verificar se houve aumento do poder preditivo em virtude da inclusão do efeito contágio.

Na primeira etapa do estudo, foi composta uma base de dados com todas as *proxies* de sentimento. As variáveis foram disponibilizadas em seu tempo t e com um *lag* $t-1$, visto que algumas incorporam alterações de comportamento de maneira célere enquanto outras apresentam atrasos, ou *lags*, na incorporação de novos choques no sentimento do mercado, conforme verificado por Baker e Wurgler (2006) e Yoshinaga (2009).

Após a composição da base de dados, realizou-se a primeira análise de componentes principais com o uso da matriz de correlação das respectivas variáveis, conforme descrito e argumentado na subseção 3.3.1, com o intuito de verificar quais os períodos, t ou $t-1$, que apresentaram maiores coeficientes no primeiro fator obtido.

Uma vez definidas as variáveis e seus respectivos períodos, foi realizada nova análise de componentes principais, desta vez apenas com cada variável em seu período de maior representatividade obtido na primeira análise. A nova análise apresentou o índice de sentimento do investidor, correspondendo à primeira etapa do trabalho proposto. Os testes de validade do índice apresentados na subseção 3.3.1 tiveram por objetivo testar a primeira hipótese proposta, indicando se houve validade estatística em relação ao índice proposto.

3 - A série histórica adotou a metodologia atual a partir do ano de 2006, porém, teve início no ano de 2003.

O mesmo procedimento utilizado na elaboração do índice de sentimento proposto, ISBR, foi adotado em um segundo índice de sentimento, ISSC, que teve excluída da sua composição a variável representativa do contágio de sentimento entre países.

Na segunda etapa deste trabalho, que teve por objetivo testar se os índices de sentimento foram significativos na predição dos retornos do ativos, adotou-se o modelo proposto por Fama e French (2015), que consiste na alocação de ativos em carteiras e fatores para verificação do poder preditivo das variáveis escolhidas nos retornos das ações, incluindo-se, neste estudo, o sentimento como variável independente.

As carteiras e os fatores foram criados com os dados obtidos no fechamento do mês de junho de cada ano, sendo os retornos verificados do mês de julho do respectivo ano até o mês de junho do ano seguinte, havendo o rebalanceamento de cada carteira para a verificação dos retornos nos doze meses seguintes. O procedimento foi repetido pelos nove anos da série.

Para a elaboração das variáveis dependentes do modelo de cinco fatores, as ações cotadas na BM&FBovespa foram inicialmente classificadas em três carteiras de acordo com o valor de mercado. Cada uma das três carteiras, sendo a primeira com as ações de baixo valor de mercado, a segunda com as ações de valor intermediário e a terceira com os maiores valores de mercado, foi então subdivida em mais três carteiras de acordo com cada uma das seguintes variáveis: *Book-to-market*, lucratividade e investimento. Da combinação do tamanho com cada uma das variáveis mencionadas foram geradas nove carteiras, totalizando vinte e sete carteiras. Os retornos dessas carteiras foram então utilizados como variáveis dependentes do modelo.

Para elaboração dos fatores, que representam as variáveis independentes utilizadas no modelo de teste, também foram obtidas as ações negociadas na BM&FBovespa, que foram então agrupadas em duas carteiras de acordo com o valor de mercado. Em seguida, as carteiras criadas foram subdivididas em duas de acordo com os índices *Book-to-market*, lucratividade e investimento, totalizando quatro carteiras para cada fator e oito carteiras no total.

Obtidas as carteiras representativas dos fatores, foram calculados os retornos mensais de cada um delas ao longo dos nove anos da amostra. Em seguida, foram subtraídos os retornos das carteiras maiores pelos retornos das carteiras menores. O resultado da subtração compôs os fatores SMB, *small minus big*, HML, *high minus low*, RMW, *robust minus weak*, e CMA, *conservative minus aggressive*. O procedimento é detalhado na subseção 3.4.2 a seguir.

Em seguida, com as carteiras e os fatores montados, foram utilizados dois instrumentos de teste, o teste GRS e regressões com dados em painel, ambos descritos na subseção 3.4.4.

Os resultados obtidos na regressão tiveram por finalidade testar a segunda hipótese proposta neste estudo, verificada através do p-valor do coeficiente angular do índice de sentimento proposto, respondendo também à pergunta de pesquisa proposta e atingindo o objetivo geral do estudo, além de confirmar ou falsear a terceira hipótese deste estudo.

A seguir estão apresentados o modelo e os procedimentos propostos para obtenção tanto do índice de sentimento quanto do seu respectivo teste.

3.3 Índice de aferição de vieses na precificação de ativos

Para a formação do índice de sentimento do investidor, tanto Baker e Wurgler (2006) quanto Yoshinaga (2009) utilizaram o primeiro fator extraído através da análise de componentes principais, descrito na subseção 3.3.1 a seguir.

Espera-se compor a variável com a estrutura proposta na equação:

$$S_t = \beta_1 NIPO_j + \beta_2 RIPO_j + \beta_3 DIVY_j + \beta_4 SNA_j + \beta_5 ICC_j \quad (31)$$

Onde S_t representa o sentimento do investidor no período t , β representa o coeficiente de cada variável, $NIPO$ representa o número de IPOs no período, $RIPO$ representa o retorno médio do primeiro dia de negociações pós-IPO, $DIVY$ representa a relação entre o *market-to-book* de pagadoras e não pagadoras de dividendos, SNA representa o sentimento do investidor norte-americano e ICC representa o índice de confiança do consumidor. Para as variáveis mencionadas, j representa a temporalidade destas variáveis, a ser apresentada e discutida em sequência, que pode assumir o período atual, t , ou o período anterior, $t-1$, de acordo com o comportamento verificado na amostra.

Uma vez proposto o modelo esperado do índice de sentimento do investidor brasileiro, apresenta-se o procedimento estatístico para a obtenção da variável pretendida.

3.3.1 Análise de componentes principais

A técnica adotada para a extração do índice de sentimento foi a análise de componentes principais, que visa “reduzir a dimensionalidade de uma base de dados que consiste em um grande número de variáveis inter-relacionadas enquanto retém o máximo

possível da variância presente na base de dados” (JOLLIFFE, 2002, p 32, tradução nossa). Abdi e Williams (2010, p. 433, tradução nossa) definem o objetivo da análise de componentes principais como o de “extrair a informação importante dos dados e expressar essa informação como um novo grupo ortogonal de variáveis chamado componentes principais”. O objetivo da análise de componentes principais se adequou ao objetivo deste trabalho, uma vez que foram extraídos índices de sentimento a partir das *proxies* propostas.

De um total composto por seis variáveis, foi extraído o primeiro fator, tendo-o como sentimento do mercado. Em consonância com o proposto, na análise de componentes principais, a partir de n variáveis, extraem-se k fatores, respeitando-se a restrição $n \ll k$. Tal redução se dá através de uma ortogonalização do sistema de coordenadas que capture a maior variância possível da amostra.

O primeiro componente é aquele que minimiza o valor de E na equação:

$$E = \varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_n^2 \quad (32)$$

Supondo que x seja um vetor com p variáveis aleatórias cuja covariância ou correlação sejam objeto de interesse por parte do pesquisador. Caso p seja um número elevado, é possível reduzir o número de variáveis ao maximizar a variância da equação apresentada a seguir (JOLLIFFE, 2002).

$$\alpha'_1 = \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1p}x_p \quad (33)$$

Onde α_1 é um vetor com p constantes $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1p}$ e α'_1 representa o vetor transposto de α_1 .

O procedimento consiste em encontrar o vetor α_1 que maximize a variância da combinação de variáveis, o que retiraria a variabilidade da amostra para determinado fator. Procedimento análogo é feito em sequência para se extrair o segundo fator, que, atendido o pressuposto de normalidade, não apresentará correlação com o primeiro dado que a amostra foi ortogonalizada em relação ao primeiro fator, podendo-se seguir com a observação dos fatores até o número p de variáveis originais.

Para encontrar o novo conjunto de coordenadas que maximize a variância de α_1 no modelo na forma matricial, tem-se:

$$\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{pmatrix} \begin{pmatrix} \alpha_{11} \\ \alpha_{12} \end{pmatrix} \quad (34)$$

Se representarmos os vetores e a matriz acima como:

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$$Z = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{pmatrix}$$

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$$

Tem-se:

$$X = Z\alpha \quad (35)$$

Onde α corresponde ao eixo do primeiro componente, cujo objetivo é encontrar os respectivos valores que agreguem a maior variância possível. Para tanto, deve-se:

$$X'X \rightarrow \text{Máx}$$

Dada a restrição abaixo em virtude da propriedade que a soma dos quadrados do seno e do cosseno de α correspondem a 1, tem-se:

$$\alpha_1^2 + \alpha_2^2 = 1 \quad (36)$$

Para resolver o problema de maximização com a restrição existente, deve-se utilizar o multiplicador de Lagrange, que resulta na equação:

$$F = \frac{\alpha'Z'Z\alpha}{n-1} - \lambda(\alpha'\alpha - 1) \quad (37)$$

Considerando que a matriz de covariância é dada por:

$$S = \frac{Z'Z}{n-1} \quad (38)$$

Tem-se:

$$F = \alpha'S\alpha - \lambda(\alpha'\alpha - 1) \quad (39)$$

Para encontrar o α que maximiza F , deve-se derivar a equação acima, que, simplificando, obtém-se:

$$S\alpha = \lambda\alpha \quad (40)$$

Sendo λ autovetor e α sendo o vetor próprio, que leva ao novo conjunto de coordenadas após o cálculo do primeiro componente. Uma vez realizada a primeira ortogonalização, parte-se para as demais ao repetir os mesmos passos com o novo conjunto de dados.

Um aspecto que deve ser considerado é a escala e a dimensão das variáveis utilizadas, pois aquelas com maiores valores absolutos levam a maiores variâncias que, por sua vez, são absorvidas ainda no primeiro fator sem necessariamente apresentarem forte relação com a variável latente em questão.

O grande inconveniente da ACP baseado na matriz de covariância é a sensibilidade dos componentes principais às unidades de medida usadas para cada elemento de x . Se houver grandes diferenças entre as variâncias dos elementos de x , então aquelas variáveis cujas variâncias são maiores tenderão a dominar o primeiro componente principal. (JOLLIFFE, 2002, p. 22, tradução nossa)

Uma forma de resolver o problema é utilizar a matriz de correlação em substituição à matriz de covariância, uma vez que mantém todas as relações dentro dos mesmos limites. Para este trabalho de natureza exploratória, utilizou-se a matriz de correlação das variáveis, uma vez que há diferentes escalas e dimensões entre estas. A questão da necessidade de normalidade da amostra não é consenso, afirmando Jolliffe (2002, p. 19) que não há tal necessidade.

Desta forma, extraiu-se o primeiro componente da amostra contendo as variáveis mencionadas na subseção 3.1.1 tomando por base a sua matriz de correlação e utilizando o fator extraído como variável de sentimento do investidor. A verificação do número de componentes principais se deu através do critério de Kaiser, que sugere a inclusão dos autovalores superiores a um.

3.4 Modelo de aferição da influência do índice no mercado

Após a obtenção da variável de sentimento, foi definido o modelo de cinco fatores proposto por Fama e French (2015) para testar se o ISBR foi significativo no retorno das ações existentes no mercado. A escolha do modelo de cinco fatores se deu em virtude dos resultados obtidos com a inclusão de dois novos fatores, que se mostraram relevantes e

contribuem para o aumento do poder de explicação dos retornos das ações negociadas no mercado. As variáveis contidas no modelo de cinco fatores mais a variável representativa do sentimento são apresentadas na equação adaptada de Fama e French (2015, p. 3):

$$R_{it} - R_{ft} = \alpha_i + \beta_i (R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + s_i S_t + e_{it} \quad (41)$$

Onde R_{it} representa o retorno da carteira i no período t , R_{ft} representa o retorno do ativo livre de risco no período t , α_i representa o intercepto, R_{mt} representa o retorno do mercado, SMB_t representa a diferença de retorno entre as carteiras compostas por pequenas e grandes empresas, HML_t representa a diferença de retorno entre as carteiras compostas por altos e baixos índices *book-to-market*, RMW_t representa a diferença entre as carteiras compostas por empresas com alta e baixa lucratividade, CMA_t representa a diferença de retorno entre as carteiras compostas por alto e baixo investimento, S representa o sentimento e, por fim, e_{it} representa o termo de erro. Os coeficientes β_i , s_i , h_i , r_i e c_i capturam as relações entre os fatores propostos e o retorno das carteiras.

3.4.1 Construção das carteiras

As variáveis dependentes adotadas por Fama e French (2015) são compostas pelo retorno das carteiras elaboradas através do agrupamento das ações de acordo com as características escolhidas: Tamanho, *book-to-market*, lucratividade e investimento.

As carteiras foram montadas com os dados obtidos no fechamento do mês de junho de cada ano. Os retornos foram calculados nos doze meses seguintes, ou seja, do mês de julho do ano t até o mês de julho do ano $t+1$, quando as carteiras foram rebalanceadas repetindo o procedimento descrito.

A partir dos dados obtidos no mês de junho de cada ano, as ações foram agrupadas em três carteiras de acordo com o seu tamanho, representado pelo valor de mercado, resultando em três carteiras iniciais, sendo a primeira representada pela letra S, *small*, a segunda representada pela letra M, *medium*, e a terceira representada pela letra B, *big*.

Cada uma das três carteiras iniciais, S, M e B, foi então subdivida em outras três carteiras para cada uma das características escolhidas. A carteira S foi novamente subdivida em três de acordo com o *book-to-market*, resultando nas carteiras SL, *small size low book-to-market*, SM_B, *small size medium book-to-market*, SH, *small size high book-to-market*. A carteira M foi dividida em três de acordo com o *book-to-market*, resultando nas carteiras ML,

medium size low book-to-market, MM_B, *medium size medium book-to-market*, MH, *medium size high book-to-market*. A carteira B foi dividida em três de acordo com o *book-to-market*, resultando nas carteiras BL, *big size low book-to-market*, BM_B, *big size medium book-to-market*, BH, *big size high book-to-market*.

De cada agrupamento foram criadas nove carteiras, resultando em vinte e sete carteiras no total, sendo cada agrupamento composto pelo tamanho e uma das três variáveis escolhidas. A figura abaixo esquematiza a elaboração das carteiras compostas pelo tamanho e pelo índice *book-to-market*:

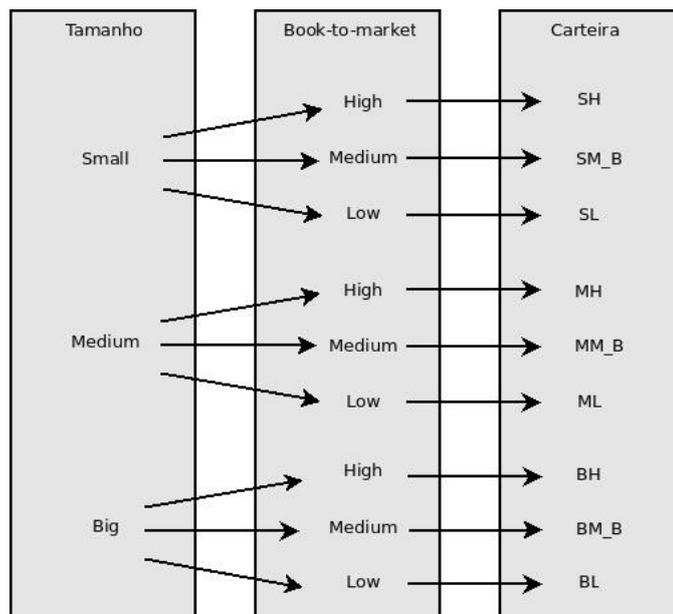


Figura 4 (3): Composição das carteiras através do tamanho e do book-to-market

Fonte: Autoria própria (2016)

Procedimento análogo foi adotado para construção das carteiras baseadas no tamanho e na lucratividade, resultando nas carteiras SW, *small size weak profitability*, SM_P, *small size medium profitability*, SR, *small size robust profitability*. A carteira M foi dividida em três de acordo com a lucratividade, resultando nas carteiras MW, *medium size weak profitability*, MM_P, *medium size medium profitability*, MR, *medium size robust profitability*. A carteira B foi dividida em três de acordo com a lucratividade, resultando nas carteiras BW, *big size weak profitability*, BM_P, *big size medium profitability*, BR, *big size robust profitability*.

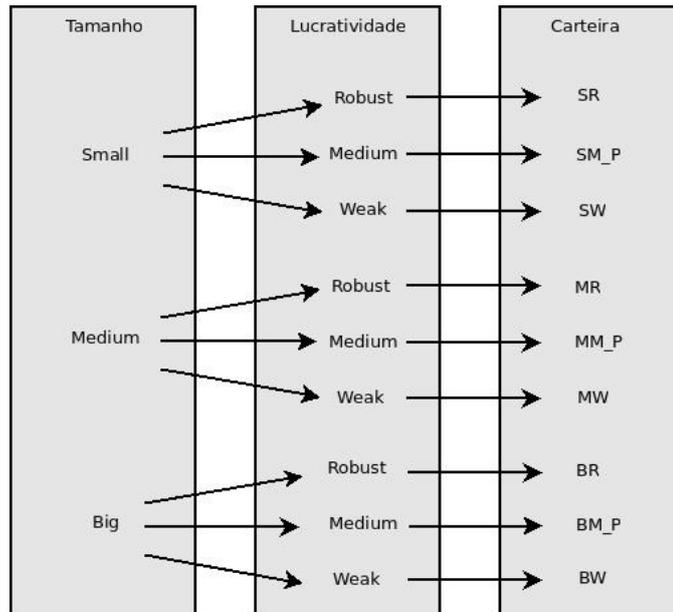


Figura 5 (3): Composição das carteiras através do tamanho e da lucratividade

Fonte: Autoria própria (2016)

Novamente foram obtidas nove carteiras compostas pelo tamanho e pela lucratividade, esquematizadas na figura acima.

Para as carteiras baseadas no tamanho e no investimento, conforme Figura 6 (3), obtiveram-se as carteiras SC, *small size conservative investment*, SM_I, *small size medium investment*, SA, *small size aggressive investment*. A carteira M foi dividida em três de acordo com o investimento, resultando nas carteiras MC, *medium size conservative investment*, MM_I, *medium size medium investment*, MA, *medium size aggressive investment*. A carteira B foi dividida em três de acordo com o investimento, resultando nas carteiras BC, *big size conservative investment*, BM_I, *big size medium investment*, BA, *big size aggressive investment*.

Em seu estudo, Fama e French (2015) dividiram as ações em cinco carteiras iniciais, seguidas por cinco subdivisões, totalizando vinte e cinco carteiras para cada conjunto de variáveis, com um total de setenta e cinco carteiras. Em virtude do menor número de ações no mercado brasileiro, optou-se por dividir as ações em três carteiras para cada variável a fim de evitar vieses em função do pequeno número de ações compondo cada carteira.

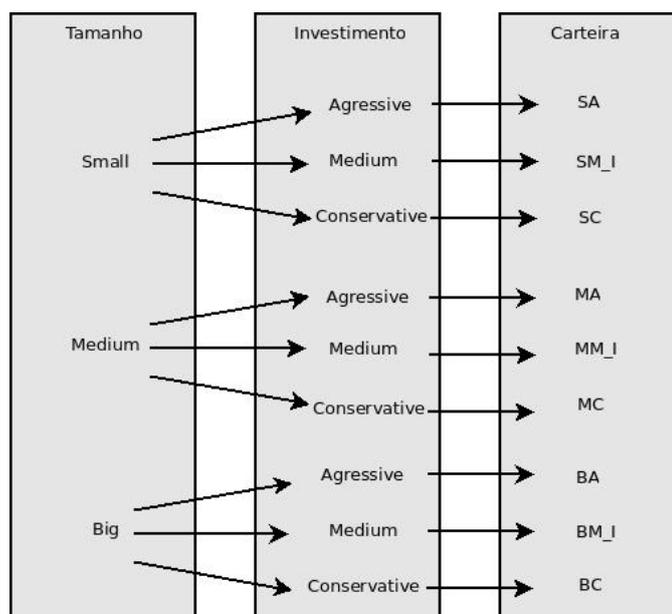


Figura 6 (3): Composição das carteiras através do tamanho e do investimento

Fonte: Autoria própria (2016)

Após a montagem das carteiras, calculou-se a média aritmética mensal do retorno das ações que as compuseram. Optou-se por ponderar igualmente os retornos em virtude da grande disparidade no valor de mercado de poucas ações, o que levou a vieses nos retornos das carteiras. Diante dos retornos médios mensais das carteiras, subtraiu-se destes o valor médio da taxa selic de cada mês, resultando no retorno médio mensal descontada a taxa livre de risco conforme proposto no modelo.

As carteiras elaboradas compuseram o conjunto de variáveis dependentes, o que os autores chamaram de LHS, *left hand side*, ou lado esquerdo da equação. A apresentação dos fatores utilizados, que compuseram as variáveis independentes do modelo, estão apresentados na subseção 3.4.2 a seguir.

3.4.2 Construção dos fatores

As variáveis independentes do modelo de cinco fatores foram compostas pelo retorno de mercado subtraído do retorno do ativo livre de risco e pelos fatores SMB, HML, RMW, e CMA descritos a seguir.

A variável representativa do retorno do mercado subtraída do retorno do ativo livre de risco foi obtida através da subtração do retorno mensal do Índice Bovespa pelo retorno da média mensal da taxa selic.

Para compor os fatores, foram inicialmente elaboradas carteiras com os dados obtidos no fechamento do mês de junho de cada ano, sendo este procedimento análogo à criação das carteiras representativas das variáveis dependentes do modelo. Os retornos das carteiras também foram calculados nos doze meses seguintes, tendo início no mês de julho do ano t e finalizando no mês de junho do ano $t+1$, quando as carteiras foram rebalanceadas e o procedimento repetido ao longo dos nove anos da série.

As ações negociadas no mercado foram inicialmente classificadas em dois grupos de acordo com o tamanho, representado pelo valor de mercado, sendo o primeiro grupo representado pelas ações de empresas com menor valor de mercado, *small*, e o segundo grupo representado pelas ações de empresas com maior valor de mercado, *big*.

A escolha da divisão em dois grupos foi em razão da sugestão dos autores quando afirmaram: “Diante das evidências aqui apresentadas, que versão dos fatores escolheríamos se começássemos de novo? Nós preferíamos os fatores a partir das carteiras 2x2 com tamanho-*book-to-market*, tamanho-lucratividade e tamanho-investimento...” (FAMA e FRENCH, 2015, p. 19, tradução nossa).

Uma vez divididas em dois grupos, *small* e *big*, as ações foram subdivididas novamente com cada um dos fatores separadamente. Desta forma, foram criadas quatro carteiras com tamanho e *book-to-market*, quatro carteiras com tamanho e lucratividade e, quatro carteiras com tamanho e investimento. O quadro abaixo apresenta a relação das variáveis que compuseram os fatores.

Quadro 3(3): Carteiras utilizadas no modelo de teste dos índices de sentimento

Tamanho	<i>Book-to-market</i>		Lucratividade		Investimento	
	<i>High</i>	<i>Low</i>	<i>Robust</i>	<i>Weak</i>	<i>Agressive</i>	<i>Conservative</i>
<i>Small</i>	SH	SL	SR	SW	SA	SC
<i>Big</i>	BH	BL	BR	BW	BA	BC

Fonte: Autoria própria, 2016.

As carteiras baseadas no tamanho e no índice *book-to-market* foram: SH, *small high*, composta pelas empresas com baixo valor de mercado e alto *book-to-market*, SL, *small low*, composta pelas empresas com baixo valor de mercado e baixo índice *book-to-market*, BH, *big high*, composta pelas empresas com alto valor de mercado e alto *book-to-market* e BL, *big low*, composta pelas empresas com alto valor de mercado e baixo *book-to-market*.

As carteiras baseadas no tamanho e no índice de lucratividade foram: SR, *small robust*, composta pelas empresas com baixo valor de mercado e alta lucratividade, SW, *small*

weak, composta pelas empresas com baixo valor de mercado e baixo índice de lucratividade, BR, *big robust*, composta pelas empresas com alto valor de mercado e alta lucratividade e BW, *big weak*, composta pelas empresas com alto valor de mercado e baixa lucratividade.

As carteiras baseadas no tamanho e investimento foram: SA, *small aggressive*, composta pelas empresas com baixo valor de mercado e alto investimento, SC, *small conservative*, composta pelas empresas com baixo valor de mercado e baixo índice investimento, BA, *big aggressive*, composta pelas empresas com alto valor de mercado e alto investimento e BC, *big conservative*, composta pelas empresas com alto valor de mercado e baixo investimento.

Após a composição das carteiras, foram calculados os retornos médios de cada uma ao longo dos doze meses seguintes, repetindo-se o procedimento de cálculo dos retornos médios e rebalanceamentos ao longo dos nove anos da série. O procedimento adotado foi análogo ao executado na construção das carteiras que compuseram as variáveis dependentes do modelo.

Após obtidos os retornos das carteiras, foram realizados os cálculos para obtenção dos valores de cada fator. O cálculo dos fatores se deu através das diferenças entre as carteiras com alto e baixo tamanho, *book-to-market*, lucratividade e investimento.

O fator HML foi obtido através da diferença entre as carteiras de alto e baixo *book-to-market*. Subtraiu-se o retorno das ações de empresas pequenas com alto *book-to-market* do retorno das ações com baixo *book-to-market* (SH - SL). Em seguida, foram subtraídos os retornos das ações de grandes empresas com alto *book-to-market* dos retornos das ações de grandes empresas com baixo *book-to-market* (BH - BL). Os resultados obtidos foram então somados e o valor obtido dividido por dois. A construção do fator HML pode ser resumida na equação (FAMA e FRENCH, 2015, p. 6):

$$HML = \frac{(SH - SL) + (BH - BL)}{2} \quad (42)$$

Analogamente ao fator HML, tem-se o cálculo do fator RMW no qual foi subtraído o retorno das ações de empresas pequenas com alta lucratividade do retorno das ações com baixa lucratividade (SR - SW). Em seguida, foram subtraídos os retornos das ações de grandes empresas com alta lucratividade dos retornos das ações de grandes empresas com baixa lucratividade (BR - BW). Os resultados obtidos foram então somados e o valor obtido dividido por dois. A construção do fator RMW pode ser resumida na equação (FAMA e FRENCH, 2015, p. 6):

$$RMW = \frac{(SR - SW) + (BR - BW)}{2} \quad (43)$$

De maneira análoga aos fatores HML e RMW, tem-se o fator CMA, representado pela diferença entre os retornos das carteiras com baixo e com alto investimento, conforme descrito na equação (FAMA e FRENCH, 2015, p. 6):

$$CMA = \frac{(SC - SA) + (BC - BA)}{2} \quad (44)$$

Para compor o fator SMB, optou-se pela diferença entre as médias de todas as carteiras com menores valores de mercado e as médias de todas as ações com maiores valores de mercado, conforme equação 45 (FAMA e FRENCH, 2015, p. 6).

$$SMB = \frac{(SH + SL + SR + SW + SC + SA)}{6} - \frac{(BH + BL + BR + BW + BC + BA)}{6} \quad (45)$$

Os fatores apresentados foram incluídos como variáveis independentes do modelo, sendo obtidos em periodicidade mensal pelos nove anos da amostra.

Uma vez obtidas as variáveis dependentes e independentes, procedeu-se com a execução do teste GRS e das regressões com dados em painel, ambos apresentados na subseção 3.4.3.

3.4.3 Teste GRS e regressões com dados em painel

A verificação da validade do ISBR como preditor dos retornos envolveu dois testes distintos. O primeiro, chamado teste GRS, foi proposto por Gibbons, Ross e Shanken (1989) e tem por objetivo verificar se os fatores utilizados para prever os retornos das carteiras são eficientes. O segundo teste envolveu o cálculo de regressões com o objetivo de verificar a validade estatística tanto dos fatores quanto dos índices de sentimento.

O teste GRS assume que, dada uma regressão múltipla, conforme equação 47, a carteira de mínima variância deve ter um intercepto α igual a zero (GIBBONS; ROSS; SHANKEN, 1989, p. 1123):

$$R_{it} = \alpha_{ip} + \beta_{ip} R_{pt} + \varepsilon_{it} \quad \forall i = 1, \dots, N, \quad (46)$$

Onde R_{it} representa o retorno em excesso do ativo i no período t , R_{pt} representa o retorno em excesso da carteira p no período t e ε_{it} o termo de erro do ativo i no período t ,

sendo o termo de erro normalmente distribuído, com média zero, variância constante e independente de R_{pt} .

A condição de mínima variância se dá porque a esperança de R_{it} deve ser atendida apenas pelo coeficiente angular multiplicado pela esperança retorno da carteira R_{pt} , conforme apresentado na equação 47 (GIBBONS; ROSS; SHANKEN, 1989, p. 1123):

$$\varepsilon(R_{it}) = \beta_{ip} \varepsilon(R_{pt}) \quad (47)$$

Desta forma, a hipótese nula, H_0 , se dá com $\alpha_{ip} = 0$ para todo i inteiro.

A não rejeição da hipótese nula do teste indica que as variáveis independentes minimizam a variância da carteira escolhida, tornando-a uma carteira eficiente em termos de média-variância e sugerindo que as variáveis independentes são suficientes para explicar o comportamento das variáveis dependentes.

Em consonância com os procedimentos adotados por Fama e French (2015), foram realizados três testes GRS para cada combinação de fatores. Cada um dos testes envolveu os mesmos fatores como variáveis independentes, porém, com apenas nove das vinte e sete carteiras como variáveis dependentes.

Do total de vinte e sete carteiras utilizadas como variáveis dependentes, ou *left hand side*, foram escolhidas as nove compostas pelo tamanho e pelo *book-to-market*, compondo assim o primeiro teste GRS, outras nove carteiras foram compostas pelo tamanho e pela lucratividade, compondo assim o segundo teste GRS, e nove pelo tamanho e pelo investimento, completando os três testes com cada grupo de fatores.

Também houve três combinações de fatores distintas. O primeiro grupo de fatores envolveu os cinco postulados pelo modelo original de Fama e French (2015), o segundo grupo de fatores foi composto pelos cinco fatores originais mais o ISBR e o terceiro grupo de fatores foi composto pelos cinco fatores mais o ISSC.

Quanto aos resultados, se a combinação do modelo de cinco fatores mais o ISBR não rejeitar a H_0 , mas rejeitá-la apenas com os cinco fatores ou com os cinco fatores mais o ISSC, conclui-se que o ISBR levou o modelo de cinco fatores à fronteira eficiente de média variância, sendo este modelo suficiente para explicar os retornos das carteiras. Tal fato confirmaria a terceira hipótese deste estudo, indicando que a inclusão da variável de sentimento melhora o poder de explicação do modelo de cinco fatores.

Caso o teste rejeite a hipótese nula apenas com o modelo de cinco fatores, mas não a rejeite com o ISBR ou com o ISSC, conclui-se que a inclusão da variável de sentimento, seja qual for, leva o modelo de cinco fatores à eficiência.

Caso haja rejeição da H_0 em qualquer situação, conclui-se que há fatores não incorporados ao modelo. Porém, mesmo com a rejeição da H_0 em todas as condições, pode-se ainda considerar as variáveis significativas, partindo-se para a análise regressiva com o intuito de verificar a validade tanto do modelo de cinco fatores quanto do ISBR e eventual superioridade deste em relação ao ISSC.

Para verificar a validade de cada índice de sentimento especificamente, optou-se pela análise dos resultados através de regressões, que tiveram por objetivo verificar se os fatores utilizados foram relevantes na determinação dos retornos das carteiras e qual o poder de explicação do modelo. Foram realizadas regressões com os cinco fatores, seguidas por regressões com os cinco fatores e os índices de sentimento ISBR e ISSC a fim de comparar eventuais diferenças no R^2 das regressões, contribuindo com a resposta à hipótese IV, sugerindo que a inclusão da variável contágio aumenta o poder de explicação do índice de sentimento sobre o retorno das carteiras.

Dado que a amostra escolhida apresenta tantos dados em corte quanto em série temporal, optou-se pela utilização de regressão com dados em painel, que reduz problemas de multicolinearidade em comparação com séries temporais e aumenta o número de graus de liberdade em comparação com dados em corte temporal (HSIAO, 2003).

Dados em painel normalmente dão ao pesquisador um grande número de dados, aumentando os graus de liberdade e reduzindo a colinearidade entre as variáveis explanatórias e conseqüentemente aumentando a eficiência das estimativas econométricas. Mais importante, dados longitudinais permitem ao pesquisador analisar um número grande de questões econômicas importantes que não poderiam ser atendidas usando dados em corte temporal ou séries temporais. (HSIAO, 2003, p. 179)

Dentre os principais tipos de regressão com dados em painel, tem-se o modelo de dados empilhados, ou *pooled data*, que consiste no agrupamento de dados em corte temporal, o modelo de efeitos fixos, ou *fixed effects*, que consiste na inclusão de variáveis *dummy* para os intervalos de tempo ou para os indivíduos, e o método de efeitos aleatórios, ou *random effects*, que consiste na inclusão de um intercepto capaz de absorver as características individuais na forma de variável aleatória. Para a escolha do modelo de dados em painel mais adequado, foram realizados os testes econométricos de Breusch-Pagan, Hausman e o estimador de efeitos fixos.

O teste de Breusch-Pagan tem como hipótese nula, H_0 , a adequação do modelo *pooled data*, sendo a hipótese alternativa, H_1 , indicativa da existência de efeitos aleatórios. O teste de Hausman, por sua vez, tem como H_0 a adequação do modelo de efeitos aleatórios, sendo a H_1 validadora do modelo de efeitos fixos. Por fim, o estimador de efeitos fixos tem como H_0 a adequação do modelo de dados empilhados, sendo a H_1 indicadora da existência de efeitos fixos.

Através dos testes acima, foi possível determinar o modelo econométrico mais adequado à amostra disponível. Uma vez realizadas as regressões, observou-se inicialmente se os coeficientes angulares das variáveis independentes foram significativos, notadamente o ISBR e o ISSC.

O foco na significância estatística do ISBR se deu em virtude deste responder a terceira hipótese deste estudo, que tem por objetivo verificar se o índice de sentimento proposto melhora o poder de explicação dos retornos das ações negociadas no mercado de capitais brasileiro.

Uma vez estatisticamente válidos, a observação da média e do erro padrão do ISSC serviu como meio de comparação com os valores obtidos no ISBR, sendo possível testar se havia diferenças significativas entre os índices, contribuindo com a elucidação do questionamento levantado na quarta hipótese deste trabalho, que afirma que a inclusão da variável representativa do contágio, representada pelo índice de sentimento norte-americano, melhora o poder de explicação do índice de sentimento do investidor brasileiro.

Também foi observado o R^2 das regressões, sendo possível verificar se houve aumento do poder de explicação com a inclusão tanto do ISBR quanto do ISSC.

3.5 Apresentação das hipóteses propostas

A estratégia metodológica escolhida tem por objetivo testar quatro hipóteses. A primeira trata da elaboração do índice de sentimento. A segunda hipótese propõe a verificação do contágio entre o sentimento brasileiro e norte-americano. A terceira hipótese trata do teste do índice de sentimento e a quarta tem por objetivo comparar o desempenho do índice com e sem a variável representativa do contágio. As quatro hipóteses são apresentadas a seguir.

Hipótese I

- Há um componente principal obtido a partir das *proxies* de sentimento observadas;

Hipótese II

- O índice de sentimento norte-americano está correlacionado com as *proxies* de sentimento do mercado brasileiro;

Hipótese III

- O índice de sentimento proposto melhora o poder de explicação dos retornos das ações negociadas no mercado de capitais brasileiro;

Hipótese IV

- A inclusão da variável representativa do contágio, representada pelo índice de sentimento norte-americano, melhora o poder de explicação do índice de sentimento do investidor brasileiro;

No intuito de apresentar os resultados e as conclusões, na seção 4 Resultados são apresentados os resultados dos procedimentos de criação dos índices e a execução dos testes propostos, incluindo a resposta à pergunta de pesquisa deste trabalho.

4 Resultados

Nesta seção são apresentados os resultados obtidos neste estudo com o objetivo de confirmar ou não as hipóteses propostas, bem como responder à pergunta de pesquisa e atingir os objetivos geral e específicos propostos.

Na subseção 4.1 são apresentados os resultados relacionados à criação do ISBR e do índice de sentimento sem a variável contágio. Esta subseção, dividida em quatro partes, oferece subsídios que permitem a confirmação das hipóteses I e II propostas. Na primeira parte, subseção 4.1.1, são apresentadas as estatísticas descritivas dos dados utilizados na elaboração dos índices de sentimento. A segunda, 4.1.2, contém os resultados da análise de componentes principais com todas as variáveis, que teve por objetivo encontrar a temporalidade adequada de cada *proxy*. Em seguida, na subseção 4.1.3, são apresentados os resultados da análise de componentes principais que gerou o ISBR. Na última, 4.1.4, são apresentados os resultados da análise de componentes principais que gerou o índice de sentimento sem a *proxy* referente ao contágio.

Em seguida, na subseção 4.2, são apresentados os resultados dos testes que visaram verificar se os índices propostos contribuíram com a explicação do retorno das ações no mercado de capitais brasileiros, levando a confirmação ou refutação das hipóteses III e IV propostas. A subseção 4.2 está dividida em cinco partes. A primeira, 4.2.1, apresenta as estatísticas descritivas das carteiras utilizadas como variáveis dependentes do modelo de testes. A segunda, 4.2.2, é composta pelos resultados dos fatores utilizados. A terceira, 4.2.3, apresenta os resultados do teste GRS. A quarta, 4.2.4, apresenta os resultados obtidos nos painéis de teste dos índices, levando à confirmação da hipótese III proposta. Por fim, na subseção 4.2.5, tem-se a comparação entre os dois índices com o objetivo de refutar ou confirmar a hipótese IV ao revelar se o ISBR apresenta superioridade em relação ao índice sem a variável contágio.

4.1 Índice de sentimento

A seleção das variáveis utilizadas na composição do índice se deu em virtude principalmente da disponibilidade de dados históricos.

A variável Contágio, que representa o índice de sentimento do investidor norte-americano, foi incluída com o intuito de encontrar evidências de contágio entre os sentimentos do investidor brasileiro e norte-americano, aspecto ainda não estudado no Brasil, mas de relevância na determinação do retorno dos ativos em virtude das evidências apresentadas por Baker Wurgler e Yuan (2012).

A variável ICC, que representa o índice de confiança do consumidor brasileiro, foi incluída em virtude de compor o corpo de *proxies* de estudos como Schmeling (2009) e Lemmon e Portniaguina (2006), o que poderia melhorar o poder preditivo do índice, além de fornecer mais evidências da validade da *proxy* mencionada para desenvolvimento de novos estudos acerca do tema.

4.1.1 Estatísticas descritivas das variáveis utilizadas nos índices de sentimento

As variáveis utilizadas na elaboração dos índices de sentimento e suas estatísticas descritivas são sintetizadas na tabela 1(4), que apresenta tanto os valores originais quanto os valores defasados em um período.

Tabela 1(4): Estatística descritiva das variáveis e suas defasagens

Estão apresentadas as estatísticas descritivas para as variáveis: número de IPOs nos últimos seis meses (NIPO), retorno médio do primeiro dia de negociação pós-IPO (RIPO), diferença entre o *book-to-market* das pagadoras menos o *book-to-market* das não pagadoras de dividendos (DIV), o sentimento do mercado norte-americano (Contágio) e o índice de confiança do consumidor (ICC). O indicativo 'd' indica que as variáveis tiveram seus valores defasados em um período.

Variável	Mínimo	1º quartil	Mediana	3º quartil	Máximo	Média	Desvio Padrão
NIPO	0	1	3,5	8	40	6,82500	8,999
RIPO	-0,0750	-0,0037	0,024100	0,06405	0,1852	0,03345	0,051
DIV	-0,6387	0,03603	0,478100	0,89690	4,5970	0,54390	0,721
Contágio	-0,8977	-0,3239	-0,19747	0,00273	0,5803	-0,1846	0,299
ICC	65	91	99	104	111	96,37	10,585
NIPO_d	0	1	4	8	40	6,867	8,979
RIPO_d	-0,075	-0,0037	0,02410	0,0645	0,18520	0,0343	0,051
DIV_d	-0,6387	0,03603	0,47810	0,8969	4,59700	0,54140	0,721
Contágio_d	-0,8977	-0,3239	-0,1915	0,01878	0,58030	-0,1826	0,299
ICC_d	65	91	99	104	111	96,54	10,264

Fonte: Autoria própria, 2016

De cada variável é apresentado o valor mínimo, o valor correspondente ao primeiro quartil, à mediana, ao terceiro quartil, ao valor máximo observado, além da média e do desvio padrão.

4.1.2 Análise de componentes principais com todas as variáveis e suas defasagens

Para a construção da série histórica que compôs o índice de sentimento, foram escolhidas as variáveis número de IPOs, retorno dos IPOs, prêmio de dividendos, sentimento norte-americano e índice de confiança do consumidor. A primeira análise de componentes principais foi composta pelas variáveis apresentadas na tabela 1(4) para que fossem obtidas as cargas fatoriais de todas as variáveis e suas defasagens. O objetivo de tal procedimento foi selecionar a temporalidade de acordo com a correlação observada no primeiro componente, sendo escolhida para compor a ACP final a que obtivesse a maior correlação com o componente. Da primeira ACP, foram obtidas as seguintes cargas:

Tabela 2(4): Cargas fatoriais da primeira análise de componentes principais

São apresentadas as cargas fatoriais do primeiro componente da análise de componentes principais para as variáveis: número de IPOs nos últimos seis meses (NIPO), retorno médio do primeiro dia de negociação pós-IPO (RIPO), diferença entre o *book-to-market* das pagadoras menos o *book-to-market* das não pagadoras de dividendos (DIV), o efeito contágio do sentimento norte-americano (Contágio) e o índice de confiança do consumidor (ICC). O indicativo 'd' indica que as variáveis tiveram seus valores defasados em um período..

Variável	1º Componente
NIPO	0,482
RIPO	0,188
DIV	0,254
Contágio	0,454
ICC	
NIPO_d	0,463
RIPO_d	0,134
DIV_d	0,159
Contágio_d	0,445
ICC_d	

Fonte: Autoria própria, 2016.

Todas as variáveis mencionadas foram agrupadas em uma base de dados tanto nos períodos originais (t) quanto com uma defasagem (t-1). O objetivo de tal agrupamento foi evidenciar a temporalidade de cada *proxy* no índice.

Conforme mencionado, o objetivo da primeira análise de componentes principais, ACP, foi observar qual temporalidade de cada variável apresentou maior correlação com o primeiro componente. Desta forma, para a obtenção do índice, foram escolhidas as variáveis NIPO, RIPO, DIV e Contágio. Observou-se também a ausência do ICC neste componente por não ter sido estatisticamente significativa em relação ao fator obtido.

4.1.3 Análise de componentes principais com o indicador de contágio.

Apesar de não estar incluído no primeiro componente da primeira ACP - Análise de Componentes Principais - o Índice de Confiança do Consumidor – ICC - foi mantido na segunda ACP para confirmação. Porém, o ICC permaneceu fora do primeiro componente mesmo sem as variáveis defasadas, sendo então retirado das análises futuras.

Em consonância com os resultados obtidos por Baker e Wurgler (2006) e Yoshinaga (2009), apenas um componente foi considerado significativo por apresentar *eigenvalue* superior a um, conforme critério de Kaiser. O *scree plot* apresentado na figura 7 (4) e a tabela 3 (4) subsequente resumem os resultados utilizados na definição do número de fatores.

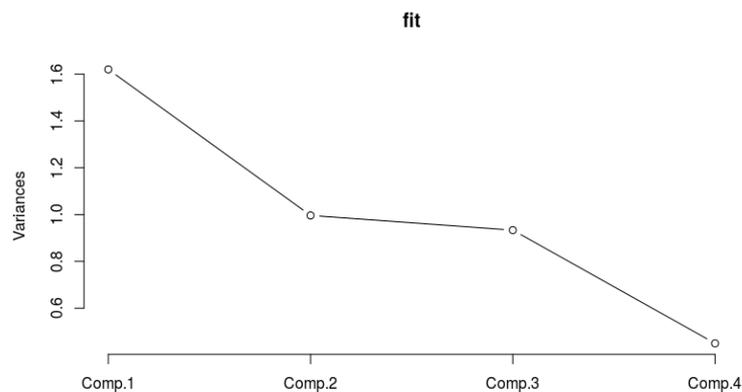


Figura 7 (4): Scree plot dos componentes da ACP

Fonte: Autoria própria fruto do *printout* do software R versão 3.3.1

A tabela 3(4) abaixo contém os valores dos componentes obtidos utilizados na composição do *scree plot* acima.

Tabela 3(4): Componentes da análise de componentes principais

São apresentados os componentes obtidos na análise de componentes principais que originou o ISBR: desvio padrão, proporção da variância explicada pelos componentes e variância acumulada pelo conjunto de componentes.

Variável	1º Componente	2º Componente	3º Componente	4º Componente
Desvio padrão	1.2729749	0.9981814	0.9662592	0.6704565
Proporção da variância	0.4051163	0.2490915	0.2334142	0.1123780
Variância acumulada	0.4051163	0.6542078	0.8876220	1.0000000

Fonte: Autoria própria, 2016.

O primeiro componente foi capaz de explicar pouco mais de quarenta por cento da variância da amostra, sendo este fator considerado representativo do sentimento do investidor. Esperou-se também que não houvesse nenhum outro fator significativo, sendo confirmada a suposição de que apenas o sentimento está relacionado às variáveis escolhidas.

Além da verificação dos componentes significativos, observou-se também as cargas fatoriais de cada variável, conforme apresentadas no quadro abaixo:

Tabela 4(4): Cargas fatoriais da segunda análise de componentes principais

São apresentadas as cargas fatoriais das variáveis representativas que compuseram o ISBR: o número de IPOs (NIPO), o retorno no primeiro dia de negociação (RIPO), a diferença entre o *book-to-market* das pagadoras e o *book-to-market* das não pagadoras de dividendos (DIV) e o índice norte-americano de sentimento (Contágio).

Variável	1º Componente
NIPO	0,676
RIPO	0,124
DIV	0,376
Contágio	0,621

Fonte: Autoria própria, 2016.

Os sinais dos valores obtidos na ACP estão de acordo com o previsto, sendo NIPO, RIPO, DIV e Contágio positivos, indicando que um aumento nesses valores levam a um aumento na variável representativa do sentimento do investidor brasileiro.

4.1.4 Análise de componentes principais sem o indicador de contágio.

Após realizada a análise de componentes principais incluindo o indicador de contágio, procedeu-se com a elaboração de nova análise sem a variável mencionada. O objetivo foi verificar se há melhoria do poder preditivo com a inclusão do índice de sentimento norte-americano.

Tabela 5(4): Componentes da ACP excluída a variável contágio

São apresentados os componentes obtidos na análise de componentes principais que originou o ISBR excluída a variável contágio: desvio padrão, proporção da variância explicada pelos componentes e variância acumulada pelo conjunto de componentes.

	1º Componente	2º Componente	3º Componente
Desvio padrão	1,1385983	1,0022305	0,8361387
Proporção da variância	0,4321354	0,3348220	0,2330427
Variância acumulada	0,4321354	0,7669573	1,0000000

Fonte: Autoria própria, 2016.

Os resultados sugerem a existência de um segundo fator significativo, representado pelo desvio padrão superior a um. Tal fator, por estar com um valor muito próximo ao limite do critério de kaiser, pode ter ocorrido em virtude de viés da amostra, que voltaria à situação observada na análise de componentes principais anterior, caso seja realizado novo teste com uma amostra superior aos dez anos escolhidos, ou pode ser fruto de algum aspecto não observado.

Através dos resultados obtidos na proporção da variância, é possível verificar o poder de explicação do componente. Com resultado de aproximadamente 0,43, é possível observar que o primeiro componente é capaz de explicar quarenta e três por cento da variância das três variáveis. Vale ressaltar que a variância explicada pelo primeiro componente representa a influência deste em cada variável utilizada, mas não um critério qualitativo para escolha do índice. Para verificar qual índice é mais adequado, deve-se testá-los enquanto variáveis independentes em um modelo de previsão de retornos.

Na tabela 6(4) estão disponibilizadas as cargas fatoriais de cada variável.

Tabela 6(4): Cargas fatoriais da análise de componentes principais sem a inclusão da variável contágio

São apresentadas as cargas fatoriais das variáveis representativas que compuseram o ISBR excluída a variável contágio: o número de IPOs (NIPO), o retorno no primeiro dia de negociação (RIPO) e a diferença entre o book-to-market das pagadoras e o book-to-market das não pagadoras de dividendos (DIV).

Variável	1º Componente
NIPO	0,710
RIPO	0,220
DIV	0,669

Fonte: Autoria própria, 2016.

Analogamente ao observado na análise de componentes principais contendo a variável Contágio, observou-se relação diretamente proporcional entre o índice de sentimento e as três variáveis.

Com os dados obtidos nos procedimentos descritos, foi possível responder às duas primeiras hipóteses propostas neste trabalho. Na primeira hipótese, que sugere a existência de um componente principal obtido a partir das *proxies* de sentimento observadas, observou-se que este componente é válido tanto no índice de sentimento proposto quanto ao ser retirada a variável Contágio, representada pelo índice de sentimento norte-americano. Desta forma, aceita-se a hipótese I como válida.

Na segunda hipótese proposta, que questionou se o índice de sentimento norte-americano está correlacionado com as *proxies* de sentimento do mercado brasileiro, observou-se a respectiva significância através da participação da variável Contágio no primeiro fator obtido, conforme pode ser verificado na tabela 4(4). Desta forma, tem-se a hipótese II como válida.

Para responder aos questionamentos propostos através das hipóteses III e IV, é apresentado a seguir o teste do índice de sentimento proposto e do seu equivalente sem a variável contágio através de um modelo de apreçamento contendo as variáveis propostas no modelo de cinco fatores de Fama e French (2015).

4.2 Aferição da influência do sentimento no retorno dos ativos

Nesta subseção são apresentados os procedimentos e os resultados obtidos através da utilização do modelo de apreçamento de cinco fatores com a adição das variáveis sentimento do investidor brasileiro, ISBR, e sentimento do investidor sem contágio, ISSC, com o intuito de testar as hipóteses III e IV, bem como responder à pergunta de pesquisa deste trabalho e atingir os objetivos geral e específicos postulados.

4.2.1 Análise descritiva das carteiras

As variáveis escolhidas para avaliar a validade do índice de sentimento proposto enquanto preditor dos retornos das ações do mercado de capitais brasileiro foram as contidas no modelo de cinco fatores proposto por Fama e French (2015). Para as variáveis dependentes, ou *left hand side*, foram criadas carteiras com dados de fechamento do mês de junho do ano t e calculados os retornos do mês de julho do ano t até o mês de junho do ano $t+1$.

Analogamente ao modelo de cinco fatores, foram criados três grupos de carteiras. O primeiro grupo composto por carteiras criadas a partir do tamanho e do índice *book-to-market*, o segundo criado a partir do tamanho e da lucratividade e o terceiro criado a partir do tamanho e do investimento.

As ações foram divididas inicialmente em três carteiras de acordo com o tamanho da empresa, representado pelo valor de mercado, em seguida subdivididas novamente em três grupos de acordo com o índice *book-to-market*, representando o primeiro agrupamento. Os retornos médios observados nos nove anos da série são apresentados na tabela 7(4).

Tabela 7(4): Estatística descritiva das carteiras Tamanho/Book-to-market

São apresentados os retornos médios das carteiras classificadas de acordo com o tamanho e o índice *book-to-market*. *Small* corresponde às carteiras com baixo valor de mercado, *medium size* corresponde às carteiras com valor de mercado intermediário e *big* correspondendo àquelas com maior valor de mercado, *high* corresponde às carteiras com maiores índices *book-to-market*, *medium BTM* corresponde às carteiras com índice *book-to-market* intermediário e *low* corresponde às carteiras com menores índices *book-to-market*.

	<i>High</i>	<i>Medium BTM</i>	<i>Low</i>
<i>Small</i>	1,114	-0,435	2,366
<i>Medium size</i>	0,076	0,564	-0,277
<i>Big</i>	-0,074	0,225	-0,256

Fonte: Autoria própria, 2016.

Observa-se que, excetuando-se a carteira de ações de baixo valor de mercado e médio *boo-to-market*, há uma tendência inversa entre o valor de mercado e o retorno das ações. Não se verifica, no entanto, relação tão evidente entre o índice *book-to-market* e o retorno das ações. Estes resultados estão de acordo com o observado por Fama e French (2015, p. 12), que questionaram a relevância do índice no modelo de cinco fatores.

O segundo agrupamento foi composto por carteiras classificadas de acordo com o tamanho e com a lucratividade, sendo apresentados os retornos médios na tabela 8(4).

Tabela 8(4): Estatística descritiva das carteiras Tamanho/Lucratividade

São apresentados os retornos médios das carteiras classificadas de acordo com o tamanho e o lucratividade. *Small* corresponde às carteiras com baixo valor de mercado, *medium size* corresponde às carteiras com valor de mercado intermediário e *big* correspondendo àquelas com maior valor de mercado, *robust* corresponde às carteiras com maiores índices de lucratividade, *medium profitability* corresponde às carteiras com lucratividade intermediária e *weak* corresponde às carteiras com menores índices de lucratividade.

	<i>Robust</i>	<i>Medium Profitability</i>	<i>Weak</i>
<i>Small</i>	1,969	0,437	0,735
<i>Medium size</i>	0,271	0,687	-0,582
<i>Big</i>	0,059	-0,010	-0,153

Fonte: Autoria própria, 2016.

Observa-se na tabela 8(4) que as carteiras com maiores índices de lucratividade obtiveram retornos superiores àquelas com menores índices de lucratividade. Um padrão inverso foi observado nas carteiras envolvendo o tamanho, no qual as carteiras *small* apresentaram retornos superiores em comparação com as carteiras *big*, apresentando uma relação inversa entre o valor de mercado e o retorno médio. No entanto, tais padrões não se mantiveram quando comparados os retornos das carteiras intermediárias com as demais.

Tabela 9(4): Estatística descritiva das carteiras Tamanho/Investimento

São apresentados os retornos médios das carteiras classificadas de acordo com o tamanho e o investimento. *Small* corresponde às carteiras com baixo valor de mercado, *medium size* corresponde às carteiras com valor de mercado intermediário e *big* correspondendo àquelas com maior valor de mercado, *agressive* corresponde às carteiras com maiores alterações no ativo total, *medium profitability* corresponde às carteiras com variação intermediária no ativo total e *conservative* corresponde às carteiras com menores alterações ou reduções no ativo total.

	<i>Agresive</i>	<i>Medium Investment</i>	<i>Conservative</i>
<i>Small</i>	0,774	0,764	1,525
<i>Medium size</i>	-0,295	0,176	0,504
<i>Big</i>	-0,349	0,017	0,234

Fonte: Autoria própria, 2016.

Pode-se observar na tabela 9(4) que tanto o valor quanto o investimento apresentaram relações inversas com o retorno. Diferentemente das demais combinações de carteiras, percebe-se que, quanto maior a alteração no valor dos ativos das empresas, menor o seu retorno. O mesmo se aplica em relação ao aumento do ativo total, que leva à redução no retorno das ações.

A carteira com maior retorno foi a SC, *small conservative*, compostas pelas empresas com menores valores de mercado e que menos aumentaram, ou mais reduziram, o ativo total. A carteira com menor retorno foi a BA, *big aggressive*, composta pelas ações das maiores empresas que mais aumentaram os seus ativos totais.

Observados os retornos médios das carteiras que representaram as variáveis dependentes do modelo, ou LHS, *left hand side*, parte-se para a análise dos fatores, que foram utilizados como variáveis independentes ou RHS, *right hand side*, do modelo de teste do índice de sentimento.

4.2.2 Análise descritiva dos fatores

Nesta subseção são apresentadas as estatísticas descritivas dos fatores utilizados como variáveis independentes do modelo de cinco fatores. Inicia-se por apresentar os resultados dos fatores $R_m - R_f$, SMB, *small minus big*, seguido pelo fator HML, *high minus low*, RMW, *robust minus weak*, e CMA, *conservative minus aggressive*.

O primeiro fator corresponde ao retorno de mercado subtraído da taxa livre de risco. A *proxy* para o retorno de mercado escolhida foi o Ibovespa, enquanto a taxa selic foi escolhida

como proxy para a taxa livre de risco. Subtraiu-se então o retorno mensal do Ibovespa da taxa selic correspondente, obtendo-se a série histórica da variável.

Cada fator restante corresponde à subtração dos retornos de duas carteiras, sendo SMB o resultado dos retornos das carteiras com alto valor de mercado subtraídos dos retornos das carteiras com baixo valor de mercado, conforme equação 46, sendo HML o resultado dos retornos das carteiras com alto *book-to-market* subtraídos dos retornos das carteiras com baixo *book-to-market*, conforme equação 43, sendo RMW o resultado dos retornos das carteiras com alta lucratividade subtraídos dos retornos das carteiras com baixa lucratividade, conforme equação 44, e CMA o resultado dos retornos das carteiras com reduções ou menores alterações nos ativos subtraídos dos retornos das carteiras com maiores altas no valor total dos ativos.

O procedimento de elaboração das séries históricas dos fatores teve início com a elaboração das carteiras. O primeiro passo foi a distribuição das ações em duas carteiras preliminares de acordo com o valor de mercado da empresa, representado pela palavra *size*.

As carteiras preliminares foram então divididas de acordo com o índice *book-to-market*, gerando assim quatro carteiras, SH, SL, BH, BL. Os valores obtidos foram então calculados conforme a equação proposta por Fama e French (2015, p. 6):

$$HML = \frac{(SH - SL) + (BH - BL)}{2} \quad (43)$$

O procedimento de cálculo foi repetido pelos cento e trinta e dois meses da série, obtendo-se a série histórica do fator HML. Os dados referentes às médias das carteiras e dos fatores estão apresentados na tabela 11(4).

Procedimento análogo foi adotado para a obtenção da série histórica do fator RMW, sendo a carteira preliminar dividida novamente de acordo com a lucratividade, gerando as carteiras SR, SW, BR e BW. Os valores obtidos no fator RMW seguiram conforme equação 44 (FAMA e FRENCH, 2015, p.6).

$$RMW = \frac{(SR - SW) + (BR - BW)}{2} \quad (44)$$

O procedimento de cálculo gerou as cento e trinta e duas observações da série histórica, estando os resultados apresentados na tabela 11(4).

Para o fator CMA, foram criadas as carteiras SC, SA, BC, BA e em seguida calculada a série histórica do fator conforme a equação 45 proposta por Fama e French (2015, p. 6).

$$CMA = \frac{(SC - SA) + (BC - BA)}{2} \quad (45)$$

Para compor o fator SMB, optou-se pela diferenças entre as médias de todas as carteiras anteriormente apresentadas com menores valores de mercado e as médias de todas as ações anteriormente apresentadas com maiores valores de mercado.

$$SMB = \frac{(SH + SL + SR + SW + SC + SA)}{6} - \frac{(BH + BL + BR + BW + BC + BA)}{6} \quad (46)$$

A tabela 10(4) apresenta os resultados dos retornos médios obtidos por cada uma das carteiras mencionadas

Tabela 10(4): Retornos médios das carteiras que compuseram os fatores

São apresentados os retornos médios das carteiras utilizadas para compor os fatores SMB, HML, RMW, CMA. SH, *small high*, composta pelas empresas com baixo valor de mercado e alto *book-to-market*, SL, *small low*, composta pelas empresas com baixo valor de mercado e baixo índice *book-to-market*, BH, *big high*, alto valor de mercado e alto *book-to-market*, BL, *big low*, alto valor de mercado e baixo *book-to-market*, SR, *small robust*, baixo valor de mercado e alta lucratividade, SW, *small weak*, baixo valor de mercado e baixo índice de lucratividade, BR, *big robust*, com alto valor de mercado e alta lucratividade e BW, *big weak*, alto valor de mercado e baixa lucratividade.

	SM	SL	BH	BL	SR	SW	BR	BW	SA	SC	BA	BC
Média	1,458	1,728	1,007	0,693	1,934	1,234	1,003	0,698	1,204	1,968	0,589	1,11

Fonte: Autoria própria, 2016.

Elaboradas as carteiras, procedeu-se com o cálculo dos fatores, cujas médias são apresentadas na tabela 11(4).

Tabela 11(4): Valores médios dos fatores HML, RMW, CMA e SMB

São apresentados os retornos médios dos fatores HML, RMW, CMA e SMB. HML corresponde à diferença de retornos das ações com alto e baixo *book-to-market*, RMW corresponde à diferença de retornos das ações com alta e baixa lucratividade, CMA corresponde à diferença entre as ações com menores aumentos e maiores aumentos no ativo total e SMB corresponde à diferença entre as carteiras com maiores e menores empresas.

	$R_m - R_f$	HML	RMW	CMA	SMB
Média	-0,300	0,022	-0,505	0,644	0,737

Fonte: Autoria própria, 2016.

Obtidos os fatores utilizados como variáveis dependentes, partiu-se para o teste das hipóteses III e IV.

4.2.3 Teste GRS

O teste GRS (GIBBONS; ROSS; SHANKEN, 1989) tem por objetivo testar se uma regressão multivariada apresenta intercepto nulo. A consequência de tal resultado sugere que os fatores utilizados são capazes de explicar o retorno obtido pelas variáveis dependentes, sendo esta combinação de fatores um modelo eficiente de previsão de retornos.

A rejeição da hipótese nula indica que o intercepto da regressão é diferente de zero, concluindo-se que a combinação de fatores é ineficiente para explicar o retorno das carteiras. Desta forma, procurou-se testar se o modelo de cinco fatores foi eficiente mesmo sem nenhuma variável de sentimento. Em seguida, incluiu-se no modelo de cinco fatores o ISBR. Por fim, incluiu-se o modelo de cinco fatores adicionado da variável ISSC.

Para realizar os testes, as carteiras que compuseram o *left hand side* do modelo foram separadas em três combinações de carteiras, seguindo os procedimentos adotados por Fama e French (2015).

Do total de vinte e sete carteiras iniciais utilizadas como variáveis dependentes, foram separadas as nove carteiras representativas do tamanho e do *book-to-market* e classificadas como primeira combinação de variáveis dependentes, a segunda combinação teve as nove carteiras envolvendo tamanho e lucratividade como variáveis dependentes e a terceira combinação teve as nove carteiras representativas do tamanho e do investimento como variáveis dependentes.

Para cada combinação de carteiras, foram realizados três testes GRS com diferentes combinações de fatores, sendo o primeiro composto apenas pelos cinco fatores originais, o segundo pelos cinco fatores mais a variável ISBR e o terceiro composto pelos cinco fatores mais a variável ISSC. Os resultados podem ser observados na tabela 12(4).

Tabela 12(4): Resultados dos testes GRS

São apresentados os p-valores dos testes GRS realizados. Foram realizados três testes tendo como variáveis dependentes as carteiras que envolviam tamanho e *book-to-market*, três testes tomando como variáveis dependentes as carteiras envolvendo tamanho e lucratividade e três envolvendo as carteiras tamanho e investimento. Para cada conjunto de variáveis dependentes foram testados três grupos de fatores, sendo o primeiro o modelo clássico de cinco fatores, o segundo incluindo também a variável ISBR e o terceiro envolvendo os cinco fatores mais o ISSC.

Carteiras	Cinco fatores	Cinco fatores + ISBR	Cinco fatores + ISSC
Tamanho-BTM	0,041	0,040	0,392
Tamanho-Lucratividade	0,133	0,137	0,136
Tamanho-Investimento	0,491	0,492	0,501

Fonte: Autoria própria, 2016.

Observa-se que apenas as carteiras compostas pelo tamanho e *book-to-market* rejeitaram a hipótese nula de eficiência em relação aos fatores. Para as variáveis compostas pelo tamanho e lucratividade e tamanho e investimento, tem-se como eficiente tanto o modelo de cinco fatores quanto o modelo de cinco fatores mais o ISBR e o modelo de cinco fatores mais o ISSC.

Os resultados indicam que os dois últimos agrupamentos de carteiras escolhidas são eficientes mesmo sem o ISBR ou sem o ISSC, não sendo possível determinar, através deste testes, se há significância nas variáveis representativas de sentimento, sendo necessário optar pela análise regressiva para confirmar ou falsear a hipótese III deste estudo.

4.2.4 Aferição dos efeitos do sentimento no retorno dos ativos

Nesta subseção são apresentados os resultados das regressões realizadas com o objetivo de verificar se as variáveis representativas do sentimento foram consideradas significativas, permitindo responder o questionamento levantado nas hipóteses III e IV.

Para cada combinação de fatores, sendo a primeira combinação composta pelo modelo de cinco fatores mais o ISBR e a segunda composta pelo modelo de cinco fatores mais o ISSC, foram realizadas três regressões.

A primeira das três regressões com cada conjunto de fatores teve como variáveis dependentes as carteiras compostas pelo tamanho e índice *book-to-market*, a segunda regressão teve como variáveis dependentes as carteiras compostas pelo tamanho e lucratividade e a terceira teve como variáveis as carteiras compostas pelo tamanho e investimento.

A combinação de variáveis dependentes, representadas pelas carteiras, e de variáveis independentes, representadas pelos fatores e índices de sentimento, totalizou seis regressões.

As regressões com o ISBR tiveram por objetivo responder ao questionamento levantado na hipótese III. A regressão com o ISSC teve por objetivo comparar os índices de sentimento, confirmando ou falseando a hipótese IV e respondendo a pergunta de pesquisa proposta.

Esta subseção tem início com os resultados dos testes que levaram à escolha do modelo de regressão mais adequado. Em seguida, são apresentados os resultados das regressões que tiveram o ISBR incluído entre os fatores, respondendo a hipótese III deste estudo, seguido pelos resultados das regressões que tiveram o ISSC incluído entre os fatores.

Por fim, são comparados os índices ISBR e ISSC, respondendo ao questionamento levantado na hipótese IV deste trabalho.

4.2.4.1 Escolha do modelo de regressão

As regressões com dados em painel podem ser realizadas através do modelo de mínimos quadrados ordinários com as séries empilhadas, ou *pooled data*, do modelo de efeitos fixos, ou *fixed effects*, ou do modelo de efeitos aleatórios, chamado *random effects*. Para verificar qual modelo de dados em painel melhor se adequou à amostra obtida, foram executados três testes: o estimador de efeitos fixos, o teste de Breusch-Pagan e o teste de Hausman. Cada um dos testes, conforme descrito a seguir, tem por objetivo comparar dois tipos de painel e indicar qual deles é mais adequado à amostra.

O primeiro teste executado foi o estimador de efeitos fixos, que tem como objetivo comparar o modelo de dados empilhados com o modelo de efeitos fixos. A hipótese nula sugere a adequação do modelo de dados empilhados, ou *pooled data*, enquanto a hipótese alternativa sugere a existência de efeitos fixos. A segunda linha da tabela 13(4) apresenta o resultado dos testes dos estimadores de efeitos fixos.

Das três regressões, apenas a terceira, representada pelas carteiras agrupadas de acordo com o tamanho e investimento, não rejeitou a hipótese nula, indicando que os dados empilhados foram mais indicados do que o painel de efeitos fixos.

Após a realização do diagnóstico de painel nas três combinações de variáveis dependentes, procedeu-se com o teste de Breusch-Pagan, que tem por hipótese nula a ideia de que o modelo com dados empilhados é mais adequado do que o modelo de efeitos aleatórios. A hipótese alternativa rejeita a adequação do MQO, indicando como mais adequado o modelo de efeitos aleatórios.

A estatística do teste Breusch-Pagan é qui-quadrado com um grau de liberdade. Novamente, apenas a terceira combinação de carteiras não rejeitou a hipótese nula, indicando que o modelo de dados empilhados foi mais adequado em comparação com o modelo de efeitos aleatórios apenas para as carteiras compostas pelo tamanho e investimento.

Com os resultados dos teste apresentados, foi possível definir que o modelo de regressão com dados empilhados, ou *pooled data*, foi mais adequado às regressões cujas carteiras foram compostas a partir do tamanho e do investimento.

Para definir o modelo mais adequado às regressões que tiveram como variáveis dependentes o tamanho e o *book-to-market* e o tamanho e a lucratividade, procedeu-se com o teste de Hausman, que tem por hipótese nula a indicação de que o modelo de efeitos aleatórios é mais adequado. A rejeição da hipótese nula sugere a adoção do modelo de efeitos fixos.

Conforme pode ser observado na tabela 13(4), o teste de Hausman rejeitou a hipótese nula tanto nas regressões cujas variáveis dependentes foram as carteiras compostas pelo tamanho e *book-to-market* quanto pelas carteiras compostas pelo tamanho e lucratividade, indicando nos dois casos que o modelo de efeitos fixos foi adequado à amostra obtida.

O resultado dos testes está apresentado na tabela 13(4) a seguir:

Tabela 13(4): Diagnóstico de Paineis

São apresentados os resultados dos testes para a escolha do modelo de regressão. Foram realizados os testes estimador de efeitos fixos, Breusch-Pagan e Hausman nas carteiras formadas a partir do tamanho e *book-to-market*, tamanho e lucratividade e tamanho e investimento. Cada teste apresenta a estatística de teste e o p-valor obtido.

Teste	Estatística	<i>Book-to-market</i>		Lucratividade		Investimento	
		Valor	P-valor	Valor	P-valor	Valor	P-valor
Estimador de efeitos fixos	F(8,957)	3,241	0,001	2,378	0,015	1,573	0,129
Breusch-Pagan	Qui-quadrado(1)	15,472	0,000	5,542	0,019	0,728	0,394
Hausman	Qui-quadrado(5)	-1,079	0,000	-3,237	0,000	-6,478	0,000

Fonte: Autoria própria, 2016.

Com base nos resultados observados, optou-se por painéis de efeitos fixos para as regressões com carteiras formadas a partir do tamanho e índice *book-to-market* e tamanho e lucratividade. Para as carteiras formadas a partir do tamanho e investimento, optou-se pela regressão com dados empilhados.

4.2.4.2 Resultados das regressões com as carteiras tamanho e *book-to-market*

A primeira regressão foi composta por painéis de efeitos fixos incluindo os cinco fatores, *dummies* temporais, mais o ISBR como variáveis dependentes e as carteiras tamanho e *book-to-market* como variáveis independentes. O modelo está representado na equação 48.

$$R_{it} - R_{ft} = \alpha_i + \beta_i (R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + D_t + ISBR_t + e_{it} \quad (48)$$

Onde R_{it} representa o retorno da carteira i formada com base no tamanho e índice *book-to-market* no período t , R_{ft} representa o retorno do ativo livre de risco, α representa o intercepto da regressão, R_{mt} representa o retorno do mercado no período t , SMB representa o

retornos das carteiras compostas por empresas com alto valor de mercado subtraídos dos retornos das carteiras compostas por empresas com baixo valor de mercado, HML representa a diferença de retornos entre as carteiras com alto e baixo *book-to-market*, RMW representa a diferença de retornos entre as carteiras com alta e baixa lucratividade, CMA representa a diferença de retornos entre as carteiras com maiores reduções e maiores aumentos no ativo total, D_t representa uma variável *dummy* para cada período e $ISBR_t$ representa o valor do índice no período t .

A segunda regressão foi composta pelo mesmo modelo, mas com o ISBR substituído pelo ISSC. As variáveis dependentes também foram os retornos das carteiras compostas pelo tamanho e pelo índice *book-to-market*, conforme equação 49.

$$R_{it} - R_{ft} = \alpha_i + \beta_i (R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + D_t + ISSC_t + e_{it} \quad (49)$$

Com a execução das regressões, foram realizados os testes de normalidade e heterocedasticidade dos resíduos. O teste de normalidade de Doornik-Hansen rejeitou a hipótese nula de normalidade dos resíduos em todas as regressões. Porém, segundo Gujarati e Porter (2011, p. 120),

[...] em amostras grandes, as estatísticas t e F têm aproximadamente as distribuições probabilísticas de t e F de forma que os testes de t e F que se baseiam na hipótese de que o erro padrão tem distribuição normal ainda possam ser aplicados validamente. [...] Portanto, a hipótese de normalidade pode não ser muito relevante em grandes conjuntos de dados.”

Os resíduos também apresentaram heterocedasticidade, verificada através do teste de Wald, apresentando p-valor igual a zero. Para corrigir o problema, foram feitos novos painéis com erros padrão robustos, o que permitiu a análise dos dados e a verificação das variáveis significativas.

Tabela 14(4): Painel de efeitos fixos com carteiras tamanho e *book-to-market*

Estão apresentados os resultados das regressões pelo método dos mínimos quadrados ordinários com painéis de efeitos fixos para as carteiras compostas pelo tamanho e *book-to-market* (Tamanho-BTM) como variável dependente e retorno de mercado ($R_m - R_f$), diferenças entre retorno das carteiras com pequenas e grandes empresas (SMB), diferenças de retorno entre carteiras com altos e baixos *book-to-market* (HML), diferenças de retorno entre carteiras com alta e baixa lucratividade (RMW), diferenças de retorno entre carteiras com alto e baixo investimento (CMA), sentimento do investidor brasileiro (ISBR) e sentimento do investidor sem contágio (ISSC) como variáveis independentes. * representa significativo ao nível de 10%, ** significativo ao nível de 5% e *** significativo ao nível de 1%.

Variável	Tamanho-BTM					
	Coefficiente	Erro Padrão	P-valor	Coefficiente	Erro Padrão	P-valor
Constante	-0,995	0,303	0,011**	0,342	-1,171	0,009***
$R_m - R_f$	1,573	0,210	0,000***	0,163	1,371	0,000***
SMB	1,883	0,269	0,000***	0,236	1,615	0,000***
HML	1,133	0,693	0,141	0,528	-0,587	0,299
RMW	3,472	1,150	0,017**	0,713	1,987	0,0237**
CMA	1,485	0,746	0,082*	0,557	0,829	0,175
ISBR	-1,674	0,582	0,021**			
ISSC				-0,309	0,889	0,021**

Fonte: Autoria própria, 2016.

As regressões realizadas conforme equações 48 e 49 com erros padrão robustos apresentaram significância estatística para todos os fatores. Os resultados são apresentados na tabela 14(4).

As duas regressões apresentaram significância no intercepto da regressão, conforme apontado anteriormente no teste GRS, no retorno de mercado, na lucratividade e nos índices de sentimento. O fator HML não foi significativo em nenhuma das regressões e o fator CMA foi apenas marginalmente significativo na regressão que teve o ISBR como fator.

Os valores observados no R^2 ajustado das regressões foram 0,62, indicando que os modelos foram capazes de explicar o retorno das carteiras em 62% aproximadamente.

Conforme pode ser observado, tanto o ISBR quanto o ISSC foram significativos ao nível de cinco por cento com ambos apresentando p-valor igual a 0,021. Os resultados contribuem para confirmar a hipótese III desta tese, mas não dão indícios sobre a validade da hipótese IV.

4.2.4.3 Resultados das regressões com as carteiras tamanho e lucratividade

Em procedimento análogo ao adotado na subseção 4.2.4.1, foram realizadas regressões tendo como variáveis dependentes os retornos das carteiras criadas com base no tamanho e na lucratividade. Para compor as variáveis independentes, foram utilizados os mesmos fatores da subseção 4.2.4.1, que foram $R_m - R_f$, SMB, HML, RMW, CMA, uma variável *dummy* para cada período e um índice de sentimento de acordo com a regressão.

A primeira regressão foi composta pelos fatores mencionados tendo o ISBR como índice de sentimento enquanto a segunda regressão foi composta pelo ISSC, ambas em painéis de efeitos fixos com erros padrão robustos em virtude de heterocedasticidade dos resíduos. Os testes de normalidade de ambas regressões rejeitaram a hipótese nula de normalidade dos resíduos.

Conforme apresentado na tabela 15(4), houve significância estatística ao nível de cinco por cento de todos os fatores na regressão composta pelo ISBR. Nas regressões compostas pelo ISSC, não foi observada significância no fator CMA, enquanto o fator HML apresentou-se apenas marginalmente significativo. Os valores observados no R^2 ajustado das regressões foram 0,64.

Tabela 15(4): Painel de efeitos fixos com carteiras tamanho e lucratividade

Estão apresentados os resultados das regressões pelo método dos mínimos quadrados ordinários com painéis de efeitos fixos para as carteiras compostas pelo tamanho e lucratividade (Tamanho-Lucratividade) como variável dependente e retorno de mercado ($R_m - R_f$), diferenças entre retorno das carteiras com pequenas e grandes empresas (SMB), diferenças de retorno entre carteiras com altos e baixos book-to-market (HML), diferenças de retorno entre carteiras com alta e baixa lucratividade (RMW), diferenças de retorno entre carteiras com alto e baixo investimento (CMA), sentimento do investidor brasileiro (ISBR) e sentimento do investidor sem contágio (ISSC) como variáveis independentes. * representa significativo ao nível de 10%, ** significativo ao nível de 5% e ***significativo ao nível de 1%.

Variável	Tamanho-Lucratividade					
	Coeficiente	Erro Padrão	P-valor	Coeficiente	Erro Padrão	P-valor
Constante	-0,977	0,382	0,034**	-1,151	0,417	0,025***
$R_m - R_f$	1,565	0,184	0,000***	1,366	0,151	0,000***
SMB	1,869	0,317	0,000***	1,605	0,298	0,001*
HML	-1,127	0,377	0,018**	-0,588	0,276	0,066**
RMW	3,391	1,134	0,017**	1,927	0,810	0,045
CMA	1,451	0,595	0,041**	0,803	0,457	0,117
ISBR	-1,652	0,418	0,004***			
ISSC				-0,877	0,222	0,004***

Fonte: Autoria própria, 2016.

A variável ISBR apresentou significância estatística ao nível de um por cento, corroborando, junto com a regressão realizada na subseção 4.2.4.1, a validade da hipótese III proposta. Também foi encontrada significância ao nível de um por cento na variável ISSC.

Observou-se que o valor obtido no coeficiente angular do ISBR, -1,652, aproximou-se do dobro do valor obtido pelo ISSC, que ficou em -0,877. Comportamento semelhante ocorreu nos erros padrão, sendo 0,418 para o ISBR e 0,222 para o ISSC. Tal diferença, no entanto, não alterou o p-valor no qual ambos obtiveram 0,004. Desta forma, não se faz possível, através da regressão apresentada na tabela 15(4), confirmar ou falsear a hipótese IV desta tese.

4.2.4.4 Resultados das regressões com as carteiras tamanho e investimento

Por fim, foram realizadas duas regressões tendo como variáveis dependentes os retornos das carteiras compostas pelo tamanho e pela alteração no ativo total das empresas. Diferentemente das demais regressões, estas tiveram o modelo de dados empilhados como mais adequado.

Tabela 16(4): Pannel de dados empilhados com carteiras tamanho e investimento

Estão apresentados os resultados das regressões pelo método dos mínimos quadrados ordinários com dados empilhados para as carteiras compostas pelo tamanho e investimento (Tamanho-Investimento) como variável dependente e retorno de mercado ($R_m - R_f$), diferenças entre retorno das carteiras com pequenas e grandes empresas (SMB), diferenças de retorno entre carteiras com altos e baixos book-to-market (HML), diferenças de retorno entre carteiras com alta e baixa lucratividade (RMW), diferenças de retorno entre carteiras com alto e baixo investimento (CMA), sentimento do investidor brasileiro (ISBR) e sentimento do investidor sem contágio (ISSC) como variáveis independentes. * representa significativo ao nível de 10%, ** significativo ao nível de 5% e ***significativo ao nível de 1%.

Variável	Tamanho-Investimento					
	Coeficiente	Erro Padrão	P-valor	Coeficiente	Erro Padrão	P-valor
Constante	-7,264	1,975	0,006***	-4,204	1,370	0,015**
$R_m - R_f$	0,147	0,113	0,000***	0,991	0,093	0,000***
SMB	0,999	0,268	0,006***	0,964	0,271	0,007***
HML	0,176	0,440	0,028**	0,969	0,440	0,058*
RMW	-3,326	1,696	0,086*	-2,501	1,612	0,159
CMA	-1,446	0,584	0,039**	-1,122	0,577	0,088*
ISBR	-5,002	1,224	0,004***			
ISSC				-1,701	0,417	0,004***

Fonte: Autoria própria, 2016.

Os resíduos das duas regressões rejeitaram a hipótese nula de homocedasticidade, sendo realizadas novas regressões com erros padrão robustos. Os testes de normalidade executados também rejeitaram a hipótese nula, seguindo o padrão observado nas demais regressões. Os valores observados no R^2 ajustado das regressões foram 0,60, indicando que os modelos foram capazes de explicar o retorno das carteiras em 60% aproximadamente.

Conforme apresentado na tabela 16(4), na regressão com o ISBR foram significativos o intercepto e os fatores $R_m - R_f$, SMB, HML, CMA e ISBR. O fator RMW apresentou significância marginal apenas.

A regressão com o ISSC como fator representativo de sentimento apresentou significância no intercepto e nos fatores $R_m - R_f$, SMB, e ISSC. Os fatores HML e CMA apresentaram significância marginal apenas. O fator RMW não apresentou significância estatística.

Com a verificação da validade estatística da variável ISBR nos três conjuntos de regressões, responde-se a terceira hipótese deste trabalho, que questiona se o ISBR melhora o poder de explicação dos retornos das ações negociadas no mercado de capitais brasileiro. Tem-se como positiva a resposta através da observação do p-valor inferior a 1% na variável mencionada.

Também é atingido o objetivo geral deste trabalho, que consistiu em verificar o efeito no retorno das ações da BM&FBOVESPA quando considerado o sentimento do mercado norte-americano na propensão à geração de vieses dos investidores brasileiros no período de 2006 a 2015, representado pelo coeficiente angular do ISBR.

Pode-se afirmar que, para o período analisado, o ISBR contribuiu com as variáveis propostas no modelo de cinco fatores na previsão dos retornos dos ativos do mercado de capitais brasileiro. Observou-se que o sentimento do investidor do mercado de capitais brasileiro sofreu influência do sentimento do investidor norte-americano e influenciou o retorno dos ativos do mercado de capitais no período considerado.

Para tentar verificar uma provável melhoria no desempenho do ISBR em virtude da variável contágio, foi realizado novo teste sem a inclusão desta no mesmo modelo de apreçamento conforme descrito na subseção 4.2.3 a seguir.

4.2.3 Comparação entre os índices de sentimento

Apresentada significância estatística no índice de sentimento do investidor brasileiro, ISBR, e no índice de sentimento sem a variável contágio, ISSC, procedeu-se com a comparação entre os dois com o objetivo de esclarecer o questionamento levantado na hipótese IV desta tese, que consistiu em verificar se a inclusão da variável representativa do contágio melhora o poder de explicação do índice de sentimento do investidor brasileiro.

O primeiro critério utilizado foi comparar o R^2 das regressões com o ISBR e com o ISSC em busca de eventuais diferenças. Porém, tanto no ISBR quanto no índice de sentimento sem a variável representativa do contágio, o indicador apresentou patamares semelhantes em seus valores, estando igualmente entre 0,62 e 0,65 nos modelos com o ISBR e o ISSC.

Dado que o poder de explicação das regressões foi semelhante, procedeu-se com a comparação entre os coeficientes angulares obtidos para verificar eventuais diferenças através do teste z, conforme equação 50.

$$Z = \frac{b_{ISBR} - b_{ISSC}}{\sqrt{EPb_{ISBR}^2 + EPb_{ISSC}^2}} \quad (50)$$

Onde *ISBR* e *ISSC* representam os índices propostos, *b* representa os coeficientes angulares obtidos nas regressões e *EP* representa o erro padrão dos referidos coeficientes.

Foram calculados os escores z dos três grupos de regressões, sendo o primeiro grupo composto pelas carteiras criadas a partir do tamanho e *book-to-market*, o segundo composto pelas carteiras criadas a partir do tamanho e lucratividade e o terceiro criado a partir das carteiras compostas pelo tamanho e investimento. Os resultados são apresentados na tabela 17(4).

Tabela 17(4): Escore z dos coeficientes angulares das variáveis de sentimento

São apresentados os resultados dos testes z para a verificação de diferenças entre os coeficientes angulares obtidos nas regressões com o ISBR e com o ISSC. Tamanho e *book-to-market*, tamanho e lucratividade e tamanho e investimento correspondem às carteiras utilizadas como variáveis dependentes. * representa significativo ao nível de 10%, ** significativo ao nível de 5% e *** significativo ao nível de 1%.

Regressão	Coeficientes		Erros padrão		Escore z	p-valor	
	ISBR	ISSC	ISBR	ISSC			
Tamanho e <i>book-to-market</i>	-1,674	-0,889	0,582	0,309	1,191	0,117	
Tamanho e lucratividade	-1,652	-0,877	0,418	0,222	1,637	0,051	*
Tamanho e investimento	-5,002	-1,701	1,224	0,417	2,553	0,005	***

Fonte: Autoria própria, 2016.

Os resultados dos testes comparativos entre os coeficientes apresentaram resultados distintos. As carteiras elaboradas a partir do tamanho e *book-to-market* não rejeitaram a hipótese nula de igualdade entre as médias com um p-valor de 0,117.

As carteiras elaboradas a partir do tamanho e lucratividade ficaram no limite da rejeição com o p-valor em 0,051, não sendo este resultado conclusivo para afirmar que o coeficiente do ISBR foi estatisticamente maior que o coeficiente do ISSC.

As carteiras compostas a partir do tamanho e investimento rejeitaram a hipótese nula de igualdade entre as médias ao nível de um por cento. Este resultado indica que a inclusão da variável contágio aumentou o poder de explicação do índice de sentimento brasileiro, o que confirmaria a hipótese IV desta tese.

A partir da análise dos dados da tabela 17(4), é possível afirmar que os valores absolutos dos coeficientes são maiores para o ISBR em comparação com o ISSC e que houve diferenças estatisticamente significativas entre os coeficientes do ISBR e do ISSC nas regressões que tiveram como variáveis dependentes os retornos das carteiras montadas a partir do tamanho e investimento.

Dado que houve significância estatística das diferenças em apenas um conjunto de carteiras, conclui-se que há indícios de confirmação da hipótese IV deste estudo, que afirma que a inclusão da variável representativa do contágio melhora o poder de explicação do índice de sentimento do investidor brasileiro.

Ao responder à quarta hipótese proposta, conclui-se que foram atingidos o objetivo geral e os objetivos específicos desta tese, além de testadas as quatro hipóteses propostas. Respondidos os questionamentos levantados ao longo desta tese, parte-se para a conclusão do trabalho na seção cinco a seguir.

5 Considerações Finais

Esta pesquisa teve como objetivo verificar o efeito no retorno das ações da BM&FBOVESPA quando considerado o sentimento do mercado norte-americano na propensão à geração de vieses dos investidores brasileiros no período de julho 2006 a junho de 2015.

Para compor a série histórica do índice de sentimento, foram coletadas cinco *proxies* sugeridas em publicações anteriores. A primeira variável foi composta pelo número de IPOs no mercado nos seis meses anteriores, a segunda foi composta pelo retorno médio observado no primeiro dia de negociação da ação pós-IPO no últimos seis meses, a terceira variável foi composta pela diferença entre os índices *book-to-market* das pagadoras e das não pagadoras de dividendos, a quarta variável foi representativa do contágio entre sentimentos de diferentes países, constatado por Baker, Wurgler e Yuan (2012), e, por fim, o índice de confiança do consumidor brasileiro.

O índice foi obtido em periodicidade mensal, totalizando cento e vinte observações. Das variáveis utilizadas, apenas o índice de sentimento do consumidor, ICC, não se mostrou significativo, sendo retirado do índice de sentimento do mercado brasileiro, ISBR. A composição do índice, ao se apresentar válida, elucidou as duas primeiras hipóteses do estudo, que foi a observação de um componente principal válido no estudo obtido através das *proxies* propostas e a correlação existente entre o ISBR e a variável contágio.

Além do índice proposto, foi também calculado um índice semelhante, mas com a variável contágio ausente, chamado ISSC, cujo objetivo foi verificar se havia diferenças entre o poder explanatório de cada índice em relação ao retorno das ações negociadas no mercado de capitais brasileiro.

Para testar a validade do índice de sentimento do investidor brasileiro enquanto capaz de melhorar o poder de explicação dos retornos das ações, foi calculado o índice GRS e realizadas regressões com dados em painel com as variáveis propostas no modelo de cinco fatores de Fama e French (2015) mais o índice de sentimento. Para as variáveis dependentes, foram criadas carteiras envolvendo o tamanho e o índice *book-to-market*, o tamanho e o índice de lucratividade e o tamanho e o índice de investimento. Para as variáveis independentes, os retornos dos fatores: $R_m - R_f$, SMB, HML, RMW, CMA mais o ISBR.

As regressões foram divididas em três grupos de acordo com o tipo de variável dependente, sendo o primeiro grupo composto pelos retornos das carteiras envolvendo o tamanho e o índice *book-to-market*, o segundo composto pelo tamanho e índice de lucratividade e terceiro pelo tamanho e índice de investimento. Cada grupo de regressões conteve uma regressão com o índice de sentimento do investidor brasileiro, ISBR, e outra regressão com o índice de sentimento sem contágio, ISSC, totalizando seis regressões.

Os resultados indicaram que a variável representativa do índice de sentimento do investidor brasileiro, ISBR, foi capaz de melhorar o poder de explicação do modelo adotado, contribuindo com um índice capaz de captar vieses de sentimento que levam a erros de precificação de ativos por parte dos investidores brasileiros e respondendo à pergunta de pesquisa proposta, que questionava qual o efeito do índice de sentimento sobre o retorno das ações da BM&FBOVESPA quando considerado o contágio do mercado norte-americano.

Para procurar evidências de que a variável contágio contribui de maneira significativa nos retornos, foram comparados os coeficientes obtidos nas regressões com o ISBR e com o ISSC. Foram encontradas evidências de superioridade do ISBR nas regressões com carteiras compostas pelo tamanho e pelo investimento.

Os resultados indicaram que os índices de sentimento foram significativos na determinação dos retornos do ativos do mercado de capitais brasileiro e que o índice de sentimento que teve em sua composição a variável representativa do contágio apresentou superioridade em relação ao índice de sentimento sem a variável representativa de contágio. Tais resultados respondem positivamente o questionamento desta tese e oferecem evidências que auxiliam na discussão acerca das evidências de contágio entre sentimentos de diferentes países.

A contribuição teórica deste trabalho envolve a consideração do contágio do sentimento norte-americano no sentimento do investidor brasileiro, que leva a vieses na precificação dos ativos. Outra contribuição é a disponibilização de um índice de sentimento capaz de melhorar o poder de explicação de modelos de apreçamento, que contribui com o aumento no poder de explicação de tais modelos.

Com relação às limitações da pesquisa, observam-se as dificuldades na obtenção de *proxies* em virtude da baixa liquidez de determinados ativos ou do pequeno histórico de determinadas séries a exemplo do índice de confiança do consumidor, que apresenta série de pouco mais de dez anos e impossibilita sua consideração em um histórico de vinte anos. Outro

aspecto relacionado à disponibilidade dos dados é o extenso período sem IPOs no Brasil no final da década de 1990 e início da década de 2000, que gera valores ausentes na proxy retorno dos IPOs e pode gerar dificuldades no procedimento estatístico de captura do índice.

Como sugestão para estudos futuros, há a inclusão de índices de diferentes países para observação de contágios ou até a utilização de um índice mundial de sentimento caso venha a ser disponibilizado e a elaboração de carteiras a partir de outras variáveis que representem as ações cuja precificação envolva mais subjetividade.

Outra alternativa envolve a utilização de diferentes procedimentos econométricos na obtenção do índice de sentimento a exemplo do *three-pass regression filter* proposto por Kelly e Pruitt (2015) ou a utilização de outras variáveis no modelo de apreciação dos ativos, que podem aumentar o poder explicativo em relação ao retorno das ações.

Referências

- ABDI, H.; WILLIAMS, L. J. Principal component analysis. **Cumputational Statistics**, v. 2, n. 4, p. 433–470, jul. 2010.
- ALAGIDEDDE, P. Return behaviour in Africa's emerging equity markets. **Quarterly Review of Economics and Finance**, v. 51, n. 2, p. 133–140, 2011.
- ALLAIS, P. M. Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. **Econometrica**, v. 21, n. 4, p. 503–546, 1953.
- BAKER, M.; WURGLER, J. Investor Sentiment and the Cross-Section of Stock Returns. **The Journal of Finance**, v. 61, n. 4, p. 1645–1680, 2006.
- BAKER, M.; WURGLER, J.; YUAN, Y. Global, local, and contagious investor sentiment. **Journal of Financial Economics**, v. 104, n. 2, p. 272–287, 2012.
- BANZ, R. W. The relationship between Return and market Value of Common Stocks. **Journal of Financial Economics**, v. 9, n. 1, p. 13–18, 1981.
- BARBERIS, N.; SHLEIFER, A.; VISHNY, R. A model of investor sentiment. **Journal of Financial Economics**, v. 49, n. 3, p. 307–343, 1998.
- BENARTZI, S.; THALER, R. H. Myopic Loss Aversion and the Equity Premium Puzzle. **The Quarterly Journal of Economics**, v. 110, n. 1, p. 73–92, 1 fev. 1995.
- BERNOULLI, D. Exposition of a new theory on the measurement of risk. **Econometrica**, v. 22, n. 1, p. 23–36, 1954.
- BHANDARI, L. C. Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence. **Journal of Finance**, v. 43, n. 2, p. 507–528, 1988.
- BHOOTRA, A.; HUR, J. The timing of 52-week high price and momentum. **Journal of Banking and Finance**, v. 37, n. 10, p. 3773–3782, 2013.
- BROWN, G. W.; CLIFF, M. T. Investor Sentiment and Asset Valuation. **The Journal of Business**, v. 78, n. 2, p. 405–440, mar. 2005.
- CARHART, M. M. On persistence in Mutual Fund Performance. **Journal of Finance**, v. 52, n. 1, p. 57–82, 1997.
- CHENG HSIAO. **Analysis of panel data**. 2. ed. Cambridge: Press syndicate of the University of Cambridge, 2003.
- CHOU, P. H.; HO, P. H.; KO, K. C. Do industries matter in explaining stock returns and asset-pricing anomalies? **Journal of Banking and Finance**, v. 36, n. 2, p. 355–370, 2012.

CHUANG, W.-I.; LEE, B.-S. An empirical evaluation of the overconfidence hypothesis. **Journal of Banking & Finance**, v. 30, n. 9, p. 2489–2515, 2006.

DANIEL, K. D.; HIRSHLEIFER, D.; SUBRAHMANYAM, A. Investor Psychology and Security Market Under- and Overreactions. **Journal of Finance**, v. 53, n. 6, p. 1839–1886, 1998.

ELLSBERG, D. Risk, Ambiguity, and the Savage Axioms. **The Quarterly Journal of Economics**, v. 75, n. 4, p. 643, nov. 1961.

EUN, C. S.; WANG, L.; XIAO, S. C. Culture and R2. **Journal of Financial Economics (JFE)**, *Forthcoming*, n. 1952, 2013.

FAMA, E. F. Efficient Capital Markets: A Review of Theory and Empirical Work. **The Journal of Finance**, v. 25, n. 2, p. 383, maio 1970.

FAMA, E. F.; FRENCH, K. R. The Cross-Section of Expected Stock Returns. **The Journal of Finance**, v. 47, n. 2, p. 427–465, 1992.

FAMA, E. F.; FRENCH, K. R. **A five-factor asset pricing model** *Journal of Financial Economics*. [s.l.] Elsevier, 2014. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0304405X14002323>>.

FAMA, E. F.; FRENCH, K. R. A five-factor asset pricing model. **Journal of Financial Economics**, v. 116, n. 1, p. 1–22, abr. 2015.

GERVAIS, S.; ODEAN, T. Learning To Be Overconfident. **The Review of Financial Studies**, v. 14, n. 1, p. 1–27, 2001.

GIBBONS, M. R.; ROSS, S. A.; SHANKEN, J. A Test of the Efficiency of a Given Portfolio. **Econometrica**, v. 57, n. 5, p. 1121–1152, set. 1989.

GRINBLATT, M.; HAN, B. Prospect theory, mental accounting, and momentum. **Journal of Financial Economics**, v. 78, n. 2, p. 311–339, 2005.

GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**. 5. ed. Porto Alegre: AMGH Editora Ltda, 2011.

HOU, T. C. T.; MCKNIGHT, P. J. An Explanation of Momentum in Canadian Stocks. **Canadian Journal of Administrative Sciences / Revue Canadienne des Sciences de l'Administration**, v. 21, n. 4, p. 334–343, 2004.

HUR, J.; SINGH, V. Reexamining momentum profits: Underreaction or overreaction to firm-specific information? **Review of Quantitative Finance and Accounting**, 2014.

HWANG, S.; RUBESAM, A. A behavioral explanation of the value anomaly based on time-varying return reversals. **Journal of Banking & Finance**, v. 37, n. 7, p. 2367–2377, 2013.

- ISRAEL, R.; MOSKOWITZ, T. J. The role of shorting, firm size, and time on market anomalies. **Journal of Financial Economics**, v. 108, n. 2, p. 275–301, 2013.
- JIANG, G.; LI, D.; LI, G. Capital investment and momentum strategies. **Review of Quantitative Finance and Accounting**, v. 39, n. 2, p. 165–188, 2012.
- JOLLIFFE, I. **Principal Component Analysis**. New York: Springer, 2002.
- KAHNEMAN, D.; TVERSKY, A. Prospect Theory: An Analysis of Decision under Risk Daniel Kahneman; Amos Tversky. **Econometrica**, v. 47, n. 2, p. 263–292, 1979.
- KELLY, B.; PRUITT, S. The three-pass regression filter: A new approach to forecasting using many predictors. **Journal of Econometrics**, v. 186, n. 2, p. 294–316, 2015.
- KIM, D. et al. Time-varying expected momentum profits. **Journal of Banking & Finance**, v. 49, p. 191–215, 2014.
- KUROV, A. Investor sentiment and the stock market's reaction to monetary policy. **Journal of Banking and Finance**, v. 34, n. 1, p. 139–149, 2010.
- LEMMON, M.; PORTNIAGUINA, E. Consumer confidence and asset prices: Some empirical evidence. **Review of Financial Studies**, v. 19, n. 4, p. 1499–1529, 2006.
- LI, Y.; YANG, L. Prospect theory, the disposition effect, and asset prices. **Journal of Financial Economics**, v. 107, n. 3, p. 715–739, 2013.
- LINTNER, J. Lintner1965a.pdf. **The Review of Economics and Statistics**, v. 47, n. 1, p. 13–37, 1965.
- LOBÃO, J. **Finanças comportamentais: Quando a economia encontra a psicologia**. 1. ed. Lisboa: [s.n.].
- LOBE, S.; RIEKS, J. Short-term market overreaction on the Frankfurt stock exchange. **Quarterly Review of Economics and Finance**, v. 51, n. 2, p. 113–123, 2011.
- MARKOWITZ, H. Portfolio selection. **The journal of finance**, v. 7, n. 1, p. 77–91, 1952.
- MENKHOFF, L. et al. Currency momentum strategies. **Journal of Financial Economics**, v. 106, n. 3, p. 660–684, 2012.
- MILLER, M.; MODIGLIANI, F. The Journal of Business,. v. 34, n. 4, p. 411–433, 1961.
- MOSSIN, J.; MOSSIN2, J. Equilibrium in a Capital Asset Market EQUILIBRIUM IN A CAPITAL ASSET MARKET'. **Source: Econometrica Econometrica**, v. 34, n. 4, p. 768–783, 1966.
- NICHOL, E.; DOWLING, M. Profitability and investment factors for UK asset pricing models. **Economics Letters**, 2014.

ROLL, R. A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. **Journal of Financial Economics**, v. 4, p. 129–176, 1977.

SCHMELING, M. Investor sentiment and stock returns: Some international evidence. **Journal of Empirical Finance**, v. 16, n. 3, p. 394–408, 2009.

SHARPE, W. F. Capital asset prices: A theory of market equilibrium under conditions of risk. **The Journal of Finance**, v. 19, n. 3, p. 425–442, 1964.

SHEFRIN, H. M.; STATMAN, M. Explaining investor preference for cash dividends. **Journal of Financial Economics**, v. 13, n. 2, p. 253–282, 1984.

SHILLER, R. J. From Efficient Markets Theory to Behavioral Finance. **Journal of Economic Perspectives**, v. 17, n. 1, p. 83–104, 2003.

THALER, R. H. Toward a positive theory of consumer choice. **Journal of Economic Behavior & Organization**, v. 1, 1980.

YOSHINAGA, C. E. **A Relação entre índice de sentimento de mercado e as taxas de retorno das ações: Uma análise com dados em painel.** [s.l.] Universidade de São Paulo, 2009.

YUAN, T.; GUPTA, R. Chinese Lunar New Year effect in Asian stock markets, 1999–2012. **The Quarterly Review of Economics and Finance**, v. 54, n. 4, p. 529–537, 2014.

APÊNDICE A – Série histórica do índice de sentimento

Mês	Valor	Mês	Valor	Mês	Valor
1	0,68	41	-1,52	81	-1,19
2	0,72	42	-0,90	82	-1,34
3	0,78	43	-0,39	83	-1,50
4	0,80	44	-0,57	84	-1,63
5	0,96	45	-1,91	85	-1,40
6	1,40	46	-1,58	86	-0,03
7	1,68	47	-1,65	87	-0,23
8	1,15	48	-1,45	88	-0,57
9	0,49	49	-1,31	89	-0,51
10	0,81	50	-1,13	90	0,10
11	0,89	51	-1,08	91	-1,92
12	1,60	52	-0,77	92	-0,96
13	1,61	53	-0,78	93	-0,37
14	2,76	54	-0,42	94	-0,39
15	3,32	55	-0,44	95	-0,44
16	3,71	56	-0,18	96	-0,48
17	3,02	57	-0,30	97	-0,85
18	3,13	58	-0,22	98	-0,68
19	3,66	59	-0,21	99	-0,60
20	3,29	60	-0,37	100	-0,92
21	2,58	61	0,08	101	-0,84
22	3,26	62	0,00	102	-0,68
23	3,30	63	0,22	103	-0,29
24	2,31	64	-0,11	104	-0,42
25	1,28	65	0,35	105	-0,23
26	1,50	66	0,41	106	-0,27
27	1,73	67	0,29	107	-0,15
28	0,58	68	0,13	108	-0,28
29	-0,21	69	-0,01	109	0,00
30	0,11	70	-0,16	110	0,12
31	0,01	71	-0,64	111	0,33
32	-0,03	72	-1,42	112	-0,22
33	-0,64	73	-1,32	113	-0,51
34	-0,23	74	-0,77	114	0,42
35	-1,66	75	-0,78	115	0,18
36	-0,97	76	-0,46	116	0,07
37	-1,42	77	-0,54	117	-0,16
38	-1,43	78	-1,11	118	0,31
39	-1,64	79	-0,84	119	0,43
40	-1,68	80	-0,80	120	-0,44