

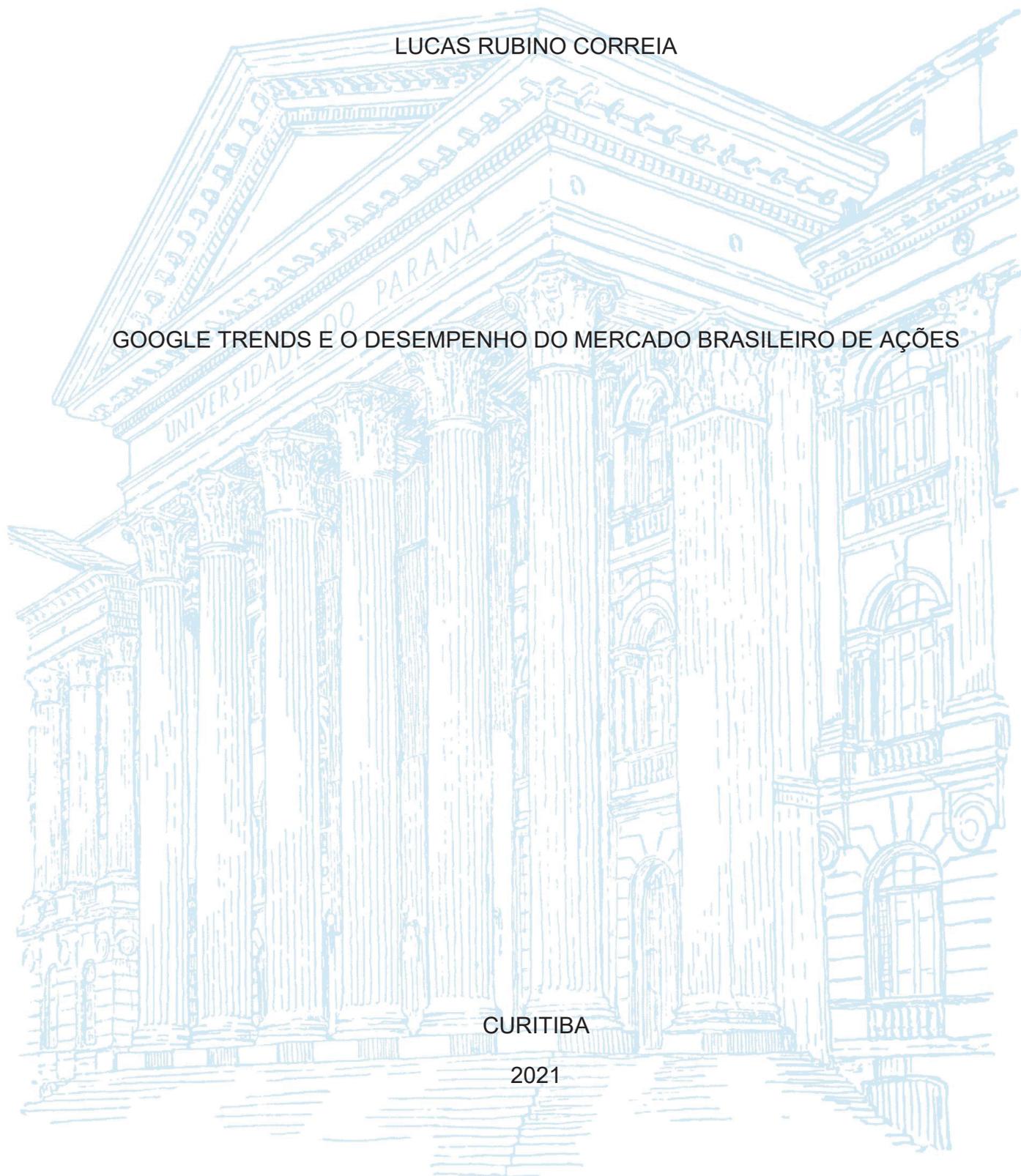
UNIVERSIDADE FEDERAL DO PARANÁ

LUCAS RUBINO CORREIA

GOOGLE TRENDS E O DESEMPENHO DO MERCADO BRASILEIRO DE AÇÕES

CURITIBA

2021



LUCAS RUBINO CORREIA

GOOGLE TRENDS E O DESEMPENHO DO MERCADO BRASILEIRO DE AÇÕES

Tese apresentada ao curso de Mestrado do Programa Profissional de Pós-Graduação em Economia, do Setor de Ciências Sociais Aplicadas da Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Economia

Orientador: Prof.^a Dr. Adalto Acir Althaus Junior

CURITIBA

2021

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DE CIÊNCIAS SOCIAIS
APLICADAS – SIBI/UFPR COM DADOS FORNECIDOS PELO(A) AUTOR(A)
Bibliotecária: Maria Lidiane Herculano Graciosa – CRB 9/2018

Correia, Lucas Rubino
GoogleTrends e desempenho do mercado brasileiro de ações / Lucas
Rubino Correia. – 2021.
123 p.

Dissertação (Mestrado) - Universidade Federal do Paraná. Programa
de Pós-Graduação em Economia, do Setor de Ciências Sociais
Aplicadas.

Orientador: Adalto Acir Althaus Junior.

Defesa: Curitiba, 2021.

1. Economia. 2. Google trends. 3. Investidores (Finanças). 4. Redes
neurais. 5. Mercado financeiro. I. Universidade Federal do Paraná.
Setor de Ciências Sociais Aplicadas. Programa de Pós-Graduação em
Economia. II. Althaus Junior, Adalto Acir. III. Título

CDD 332.6



MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS SOCIAIS E APLICADAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO ECONOMIA -
40001018051P7

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação ECONOMIA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de LUCAS RUBINO CORREIA intitulada: *GoogleTrends e Desempenho do Mercado Brasileiro de Ações*, sob orientação do Prof. Dr. ADALTO ACIR ALTHAUS JUNIOR, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 09 de Setembro de 2021.

Assinatura Eletrônica

10/09/2021 20:03:19.0

ADALTO ACIR ALTHAUS JUNIOR

Presidente da Banca Examinadora

Assinatura Eletrônica

10/09/2021 17:12:29.0

RICARDO LOBATO TORRES

Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ)

Assinatura Eletrônica

10/09/2021 17:29:13.0

RODOLFO COELHO PRATES

Avaliador Interno

AGRADECIMENTOS

Agradeço à toda minha família, em especial aos meus pais, Rubens e Isabel, por toda a educação, apoio e incentivo que vocês sempre me deram ao longo de minha caminhada pessoal e profissional.

Agradeço à minha namorada, Tiemi Tomonaga, pelo companheirismo, amor e pelos incentivos nos momentos mais necessários.

Agradeço à Universidade Federal do Paraná e ao Programa de Pós-Graduação em Economia (PPGEcon), pela oportunidade de realização deste trabalho. Agradeço aos meus colegas e amigos de mestrado e aos professores que compartilharam comigo momentos de discussões e aprendizado.

Agradeço ao meu orientador, Prof. Dr. Adalto Acir Althaus Junior, pela paciência, pelo conhecimento e pelo auxílio durante a elaboração dessa pesquisa.

Por fim, gostaria de agradecer a todos que, de alguma forma, contribuíram para a realização desse trabalho.

RESUMO

O trabalho proposto teve como objetivo avaliar se as evoluções de buscas de palavras-chave no Google, ao serem utilizadas como métricas de atenção e sentimento dos investidores, conseguem fornecer tendências de desempenho do mercado financeiro brasileiro. Com base em dados do período entre 2015 e 2019, estimamos regressões com dados em painéis e regressões VAR, para investigar o efeito da frequência de pesquisas em palavras relacionadas a sentimento financeiro, nomes de índices de mercado, *tickers* de negociação e nome de empresas sob três indicadores de desempenho de mercado: retorno semanal, volume de negociação semanal e volatilidade. Os resultados confirmam que o Google pode ser utilizado como uma ferramenta de auxílio nas previsões de desempenho de mercado, por exemplo: (i) um aumento na atenção anormal na palavra “Ibovespa” indica um impacto positivo no retorno semanal dos índices de ação Ibovespa de duas semanas posteriores; (ii) um maior nível de pesquisas em palavras que formam o índice de sentimento negativo reflete uma piora no retorno dos principais índices acionários na semana posterior. Por fim, desenvolvemos uma estrutura de rede neural artificial no software Matlab combinada com dados extraídos do Google Trends, para sugerir duas estratégias de alocação de capital e comparar as rentabilidades com o retorno do Ibovespa no mesmo período, verificamos que ambas as estratégias alcançaram rentabilidades superiores ao do índice.

Palavras-chave: Google Trends; Ibovespa; atenção dos investidores; sentimento dos investidores; redes neurais.

ABSTRACT

The purpose of the study was to evaluate if Google search queries can provide insights of the Brazilian financial market performance. Based on weekly data from the period between 2015 and 2019, we estimated regressions with panel data and VAR regressions to investigate the effect of frequency of searches on words related to financial sentiments, markets index names, trading tickers and company names over three indicators of market performance: weekly return, weekly trading volume, and volatility. The findings support that Google can be used as a tool to predict market performance, for example: (i) An increase in abnormal attention to the word "Ibovespa" indicates a positive impact on the weekly return of Ibovespa index in two subsequent weeks; (ii) a higher level of searches over words that form the negative sentiment index reflects a negative return of market index in the subsequent week. Finally, we developed an artificial neural network structure in Matlab software combined with data extracted from Google Trends, to suggest two capital allocation strategies and to compare with Ibovespa's return in the same period, we verified that both strategies achieved higher returns than the index.

Key words: Google Trends; Ibovespa; investor attention; investor sentiment; neural networks.

LISTA DE FIGURAS

| | |
|--|-----|
| FIGURA 1 - Diagrama do efeito da atenção na incorporação das informações nos preços dos ativos. | 24 |
| FIGURA 2 - Evolução do índice de pesquisas por ITUB4 no período entre 2015 e 2019. | 46 |
| FIGURA 3 - Métodos utilizados para testar cada uma das hipóteses. | 56 |
| FIGURA 4 – Gráficos das séries temporais das variáveis analisadas nessa seção.. | 60 |
| FIGURA 5 – Gráficos das séries temporais das variáveis analisadas nessa seção.. | 76 |
| FIGURA 6 – Estrutura de uma Rede Neural com várias camadas..... | 98 |
| FIGURA 7 – Script da rede neural no software MATLAB® | 100 |

LISTA DE QUADROS

| | |
|---|-----|
| QUADRO 1 – Empresas selecionadas na amostra via índice IBRX-100 | 41 |
| QUADRO 2 – Resumo das principais métricas utilizadas | 44 |
| QUADRO 3 – Termos de busca analisados por categoria. | 47 |
| QUADRO 4 – Variável conforme categoria do termo de busca..... | 49 |
| QUADRO 5 – Datas de atualização do agrupamento de palavras..... | 54 |
| QUADRO 6 – Métodos utilizados para testar cada uma das hipóteses. | 55 |
| QUADRO 7 – Características do conjunto de dados para análise. | 99 |
| QUADRO 8 – Resultado do teste da melhor configuração de rede. | 103 |
| QUADRO 9 - Características das 05 melhores configurações de redes | 104 |
| QUADRO 10 – Resultado simulação de investimentos 1 | 106 |
| QUADRO 11 – Resultado simulação de investimentos 2..... | 108 |

LISTA DE TABELAS

| | |
|---|-----|
| TABELA 1 – Índices de ações..... | 43 |
| TABELA 2 – Estatísticas descritivas – sentimento FEARS e os indicadores de mercado | 58 |
| TABELA 3 – Matriz de correlações – sentimento FEARS e os indicadores de mercado | 58 |
| TABELA 4 – Teste da raiz unitária ADF – Regressão sentimento FEARS e indicadores de mercado. | 59 |
| TABELA 5 – Testes de defasagens – AIC, SC e HQ. | 61 |
| TABELA 6 – Teste de causalidade de Granger – FEARS e retorno semanal dos índices de mercado | 64 |
| TABELA 7 – Estimativa VAR – Índices FEARS e retorno semanal dos índices de mercado | 65 |
| TABELA 8 – Teste de causalidade de Granger – índices FEARS e a volatilidade semanal do Ibovespa | 67 |
| TABELA 9 – Estimativa VAR - FEARS e a volatilidade semanal | 68 |
| TABELA 10 – Teste de causalidade de Granger – FEARS e o volume semanal de negociações | 70 |
| TABELA 11 – Estimativa VAR - FEARS e o volume semanal de negociações..... | 71 |
| TABELA 12 – Estatísticas descritivas – Atenção dos investidores e os indicadores de mercado | 74 |
| TABELA 13 – Matriz de correlações – Atenção dos investidores e os indicadores de mercado | 74 |
| TABELA 14 – Teste da raiz unitária – Atenção dos investidores e os indicadores de mercado. | 75 |
| TABELA 15 – Testes de defasagens – AIC, SC e HQ. | 77 |
| TABELA 16 – Teste de causalidade de Granger – Atenção em índices e o retorno semanal..... | 79 |
| TABELA 17 – Estimativa VAR - Atenção em índices e o retorno semanal | 80 |
| TABELA 18 – Teste de causalidade de Granger – Atenção em índices e a volatilidade semanal..... | 81 |
| TABELA 19 – Estimativa VAR - Atenção em índices e a volatilidade semanal | 82 |
| TABELA 20 – Teste de causalidade de Granger – Atenção em índices e o volume semanal e anormal..... | 84 |
| TABELA 21 – Estimativa VAR - Atenção em índices e o volume semanal e anormal | 85 |
| TABELA 22 – Estatísticas descritivas – Dados em painéis | 87 |
| TABELA 23 – Matriz de Correlação – Dados em painéis | 87 |
| TABELA 24 – Dados em painéis - Atenção em empresas/tickers e retornos anormais | 89 |
| TABELA 25 – Dados em painéis - Atenção em empresas/tickers e retornos semanais | 91 |
| TABELA 26 – Dados em painéis - Atenção em empresas/tickers e volumes anormais | 93 |
| TABELA 27 – Dados em painéis - Atenção em empresas/tickers e volumes semanais | 95 |
| TABELA 28 – Desempenho das 05 melhores configurações de rede..... | 102 |

SUMÁRIO

| | | |
|----------|--|-----------|
| 1 | INTRODUÇÃO | 13 |
| 1.1 | OBJETIVO GERAL | 15 |
| 1.2 | OBJETIVOS ESPECÍFICOS | 15 |
| 1.3 | HIPÓTESES E CONTRIBUIÇÕES | 16 |
| 2 | REVISÃO DE LITERATURA | 17 |
| 2.1 | TEORIA TRADICIONAL DAS FINANÇAS E TEORIA COMPORTAMENTAL | 17 |
| 2.2 | ATENÇÃO DO INVESTIDOR | 22 |
| 2.3 | SENTIMENTO DO INVESTIDOR | 26 |
| 2.4 | GOOGLE TRENDS E A ATENÇÃO E SENTIMENTO DOS INVESTIDORES | 29 |
| 2.5 | GOOGLE TRENDS E O MERCADO FINANCEIRO BRASILEIRO | 37 |
| 3 | METODOLOGIA | 40 |
| 3.1 | ESCOPO DO TRABALHO E DADOS | 40 |
| 3.1.1 | Ações listadas na B3 | 40 |
| 3.1.2 | Índices de mercado | 43 |
| 3.1.3 | Palavras relacionadas ao sentimento dos investidores | 43 |
| 3.2 | MÉTRICAS E COLETA DOS DADOS | 44 |
| 3.2.1 | Frequência de buscas no Google | 45 |
| 3.2.2 | Indicadores financeiros | 49 |
| 3.2.2.1 | Retornos anormais e semanais | 49 |
| 3.2.2.2 | Volumes anormais e semanais | 51 |
| 3.2.2.3 | Volatilidade semanal dos ativos | 52 |
| 3.2.3 | Construção do índice de sentimento “FEARS” | 53 |
| 3.3 | MÉTODOS E PROCEDIMENTOS ECONOMETRICOS | 54 |
| 4 | ÍNDICES FEARS E OS INDICADORES DE MERCADO | 57 |
| 4.1 | ÍNDICES FEARS E O INDICADOR DE RETORNO SEMANAL | 63 |
| 4.2 | ÍNDICES FEARS E O INDICADOR DE VOLATILIDADE SEMANAL | 67 |
| 4.3 | ÍNDICES FEARS E O INDICADOR DE VOLUME SEMANAL | 69 |
| 5 | ATENÇÃO EM ÍNDICES DE AÇÕES E OS INDICADORES DE MERCADO | 73 |
| 5.1 | ATENÇÃO EM ÍNDICES DE AÇÕES E O RETORNO SEMANAL | 78 |
| 5.2 | ATENÇÃO EM ÍNDICES DE AÇÕES E A VOLATILIDADE SEMANAL | 81 |
| 5.3 | ATENÇÃO EM ÍNDICES DE AÇÕES E O VOLUME SEMANAL/ANORMAL 83 | |
| 6 | ATENÇÃO EM TICKERS/EMPRESAS E OS INDICADORES DE MERCADO | 86 |
| 6.1 | ATENÇÃO EM TICKERS/EMPRESAS E OS RETORNOS ANORMAIS | 88 |
| 6.2 | ATENÇÃO EM TICKERS/EMPRESAS E OS RETORNOS SEMANAIS | 90 |
| 6.3 | ATENÇÃO EM TICKERS/EMPRESAS E O VOLUME ANORMAL | 92 |
| 6.4 | ATENÇÃO EM TICKERS/EMPRESAS E O VOLUME SEMANAL | 94 |
| 7 | ESTRATÉGIAS DE INVESTIMENTOS | 97 |
| 7.1 | ESTRUTURAÇÃO REDE NEURAL ARTIFICIAL (RNA) | 97 |
| 7.2 | PARÂMETROS DA REDE NEURAL ARTIFICIAL | 98 |
| 7.3 | DESENVOLVIMENTO DA REDE NEURAL ARTIFICIAL NO MATLAB | 100 |
| 7.4 | ANÁLISE DOS RESULTADOS | 102 |

| | | |
|----------|--|------------|
| 7.5 | SIMULAÇÃO DE INVESTIMENTO | 105 |
| 7.5.1 | Simulação de investimento 1 | 105 |
| 7.5.2 | Simulação do investimento 2 | 107 |
| 8 | CONCLUSÃO | 109 |
| 8.1 | LIMITAÇÕES E SUGESTÕES DE TRABALHOS FUTUROS | 111 |
| | REFERÊNCIAS..... | 112 |
| | APÊNDICE A - PALAVRAS DE SENTIMENTO DE DA ET. AL (2015) | 120 |
| | APÊNDICE B - SELEÇÃO DE PALAVRAS DE SENTIMENTOS..... | 123 |

1 INTRODUÇÃO

A atenção e o sentimento dos indivíduos são pressupostos que desafiam as ideias centrais da Hipótese do Mercado Eficiente (HME), construída com as premissas de que os investidores são sempre racionais, têm recursos cognitivos sem restrições e são capazes de dar atenção adequada a todos os tipos de informações relacionadas ao mercado. De acordo com a HME, conforme Fama (1970), as informações disponíveis para a tomada de decisão dos investidores já estariam incorporadas aos preços dos ativos, tornando irrelevante estudos que buscam a previsibilidade de movimento dos mercados.

Contrariando a HME, Kahneman (1973) afirmou que a atenção, condição necessária para absorver as informações disponíveis, é um recurso cognitivo limitado que influencia a racionalidade da tomada de decisão e impede a maximização da utilidade por parte dos investidores. Complementando, De Long et al. (1990) sugeriram que os investidores baseiam suas negociações em seus sentimentos e ruídos de mercado, gerando um excesso de volatilidade nos mercados, afastando o preço dos títulos de seus valores fundamentais e limitando a atuação dos arbitradores de Mercado.

Nesse contexto surge uma nova forma de medir a atenção e sentimento dos investidores do mercado financeiro, por meio da análise da evolução do índice de buscas no Google (GSVI, do inglês Google Search Volume Index) fornecida pela ferramenta Google Trends (GT). Os sites de pesquisas tornaram-se uma recorrente fonte de informação na internet, por satisfazerem de forma rápida e eficiente as intenções do utilizador, conforme apontado por Latoeiro (2012). Em 2019, por exemplo, o Google possuía cerca de 92% do *share* do mercado de ferramentas de pesquisas e teve sua página principal visitada 62 bilhões de vezes no ano, conforme informações retiradas dos sites Gs.statcounter (2019) e Similarweb (2019), respectivamente.

O Google Trends possibilita a geração de uma alta quantidade de dados em tempo real que facilitam a identificação e a compreensão de tendências de comportamento, do sentimento da população sobre determinados temas e o entendimento de fenômenos pontuais, como proliferação de doenças e aumento de

demanda por determinado produto. Conforme Da et al. (2011), se alguém estiver pesquisando algo no Google, está obviamente prestando atenção no tema. No âmbito das finanças, Bijl et al. (2016), Da et al. (2011), Vozlyublennaia (2014), Aoadi et al. (2013) e Preis et al. (2013), por exemplo, validaram as buscas no Google como direcionador relevante da atenção e do sentimento dos investidores.

Isto posto, o presente trabalho analisa a relação entre o comportamento dos investidores, medido via evolução de pesquisas no Google, e um conjunto de ações listadas na B3 e os principais índices de *benchmarks* de mercado, com o objetivo de verificar se as buscas de palavras-chave no Google conseguem identificar tendências de desempenho do mercado acionário.

Para estudarmos o problema e analisarmos as hipóteses sugeridas, selecionamos os índices de evoluções semanais de dados de busca no Google de palavras alusivas aos sentimentos dos investidores e consumidores, nomes de índices de mercado de ações, *tickers* de negociação de empresas e nomes de empresas negociadas no índice IBrX. Para medir os efeitos, esses dados foram relacionados, por meio de dados em painéis e regressões VAR, com três indicadores de desempenho: retorno semanal, volume de negociação semanal e volatilidade.

As análises econométricas confirmam que o Google pode ser utilizado como uma ferramenta de auxílio nas previsões de desempenho de mercado, tornando-se uma métrica para medição da atenção e do sentimento dos investidores. Dentre os resultados identificados, destacamos: (i) um aumento na atenção anormal na palavra “Ibovespa” pelos usuários do Google indica um impacto positivo no retorno semanal dos índices de ação Ibovespa de duas semanas posteriores; (ii) um maior nível de pesquisas em palavras que formam o índice de sentimento negativo (FEARS negativo) reflete uma piora no retorno dos principais índices acionários na semana posterior.

Com o objetivo de expor o potencial preditivo dessas informações, duas estratégias de compra e venda de ações utilizando dados do Google Trends foram testadas, via redes neurais, e os resultados foram comparados com o desempenho do índice Ibovespa no mesmo período analisado. Em ambas as estratégias utilizadas, a rentabilidade foi superior ao desempenho do índice.

O estudo está organizado em oito partes. A primeira é formada pela introdução, objetivos e hipóteses. A segunda apresenta uma revisão da bibliografia

sobre o tema aplicado. A terceira traz a metodologia utilizada e a base de dados. A quarta, quinta e sexta parte discute os resultados obtidos nas regressões. A sétima detalha a construção e resultados das estratégias de compra e venda de ações sugeridas e a oitava parte apresenta as conclusões.

1.1 OBJETIVO GERAL

Avaliar a relação entre a evolução de índices de buscas de palavras-chave no Google e o desempenho do mercado brasileiro de ações.

1.2 OBJETIVOS ESPECÍFICOS

- Analisar a relação entre a atenção dos investidores em determinados ativos e o desempenho de seus respectivos índices financeiros (volatilidade, volume e retorno).
- Analisar a relação entre o sentimento dos investidores, medido via evolução de buscas de palavras no Google Trends, com o desempenho do índice de mercado e de seus principais componentes.
- Sugerir possíveis estratégias de compra e venda de ações baseadas nos dados extraídos do Google Trends e na utilização de redes neurais artificiais.

1.3 HIPÓTESES E CONTRIBUIÇÕES

Visando atingir os objetivos propostos e contribuir com a literatura, este trabalho examina as seguintes hipóteses:

H1a: Um aumento nos índices de sentimento dos investidores reflete um efeito positivo ou negativo no retorno semanal dos índices de mercado.

H1b: Um aumento nos índices de sentimento dos investidores reflete um efeito positivo ou negativo na volatilidade dos índices de mercado.

H1c: Um aumento nos índices de sentimento dos investidores reflete um efeito positivo ou negativo no volume negociado dos índices de mercado.

H2a: Um aumento na atenção dos investidores nos principais índices do mercado brasileiro reflete uma mudança positiva ou negativo no retorno dos índices.

H2b: Um aumento na atenção dos investidores nos principais índices do mercado brasileiro reflete uma mudança positiva ou negativo na volatilidade dos índices.

H2c: Um aumento na atenção dos investidores nos principais índices do mercado brasileiro reflete uma mudança positiva ou negativo no volume negociado dos índices.

H3a: Um aumento na atenção dos investidores nas ações de empresas listadas na BM&F, medido via *Tickers*, reflete um efeito positivo ou negativo no retorno semanal das ações.

H3b: Um aumento na atenção dos investidores nas ações de empresas listadas na BM&F, medido via *Tickers*, reflete um efeito positivo ou negativo no volume negociado semanal das ações.

H4a: Um aumento na atenção dos investidores nas ações de empresas listadas na BM&F, medido via nome da empresa, reflete um efeito positivo no retorno semanal das ações.

H4b: Um aumento na atenção dos investidores nas ações de empresas listadas na BM&F, medido via nome da empresa, reflete um efeito positivo no volume negociado semanal das ações.

2 REVISÃO DE LITERATURA

2.1 TEORIA TRADICIONAL DAS FINANÇAS E TEORIA COMPORTAMENTAL

Economistas como Adam Smith, Irving Fisher e John Maynard Keynes argumentaram que a racionalidade imperfeita dos indivíduos afeta as decisões de investimento e o comportamento dos mercados financeiros. No entanto, por volta da década de 1950, as teorias financeiras avançaram para um paradigma de racionalidade perfeita e mercados informacionalmente eficientes. (HIRSHLEIFER; TEOH, 2003).

A teoria moderna das finanças começou a ganhar relevância com o lançamento do artigo *Portfolio Selection*, de Harry Markowitz. Markowitz (1952) assumiu que todos os investidores são racionais e lidam com duas variáveis ao aplicarem seus recursos: o retorno esperado e a variância dos retornos esperados. Ao trabalhar com essas características, o investidor busca maximizar a utilidade e minimizar o risco indesejado. A partir dos conceitos aplicados por Markowitz (1952), importantes estudos marcaram o desenvolvimento teórico das finanças, tais como: o modelo de precificação de ativos (CAPM) de Sharpe (1964) e Litner (1965), o modelo de precificação de opções de Black e Scholes (1973) e a Hipótese dos Mercados Eficientes (HME) de Fama (1970).

Proposta, inicialmente, por Roberts (1967) e aperfeiçoada por Fama (1970), a hipótese do mercado eficiente afirma que um mercado é considerado eficiente quando qualquer nova informação é incorporada imediatamente ao seu preço de negociação, não existindo espaço para a obtenção de retornos anormais e acima da média de mercado. De acordo com Fama (1991), um mercado eficiente pode ser definido como aquele composto por inúmeros agentes racionais que competem entre si e buscam prever o valor futuro do mercado, esses agentes contam com todas as informações disponíveis próximas a um custo zero. Nesse cenário, o preço de um ativo sempre será uma boa estimativa do seu valor intrínseco e, com isso, conforme Brealey e Myers (1992), as transações de compra e venda ao preço vigente nunca teriam valor presente líquido positivo.

Complementando, as condições teóricas básicas para que o mercado seja eficiente, segundo Fama (1970) são: (i) todos os participantes concordam com os efeitos das informações no atual preço de cada título, (ii) não há custos de transação nas negociações de ativos, (iii) todas as informações disponíveis devem estar ao alcance de todos os participantes sem custos de obtenção.

De acordo com Malkiel (2003), a HME está associada com a ideia de “Random Walk”, termo utilizado para explicar que os preços dos ativos evoluem de acordo com uma aleatoriedade, impedindo a previsibilidade de seus movimentos. A lógica desse movimento de mercado é que as variações de preços futuros refletem apenas novas informações futuras e serão independentes da variação de preço atual. Dessa forma, as técnicas de análise gráfica não permitiriam ao investidor obter retornos superiores aos obtidos em uma carteira selecionada aleatoriamente de ações individuais, pelo menos não com risco comparável.

Yoshinaga (2009 p.25) sugeriu que na visão de muitos estudiosos, a HME prevalece mesmo em um ambiente onde parte dos agentes atue de forma irracional, pois “ (i) se houver um desvio dos preços dos ativos em relação a seus valores corretos, cria-se uma oportunidade atrativa de lucros; (ii) agentes racionais aproveitarão rapidamente uma oportunidade, conduzindo os preços de volta aos seus valores de equilíbrio”. Essa atuação contraria o pressuposto da análise fundamentalista, que tem como objetivo avaliar, por meio dos fundamentos da empresa, variações entre o preço do ativo e seu valor intrínseco.

Nesse sentido, Damodaran (1997) afirmou que a eficiência de mercado ocorre quando o preço de mercado do título for uma estimativa não tendenciosa do valor real do ativo, onde: (i) os preços podem divergir do seu valor real desde que sejam desvios aleatórios, (ii) existe a possibilidade de que uma ação esteja sub ou superavaliada em qualquer momento, porém esses desvios não estão correlacionados com qualquer variável observável, (iii) nesse cenário, nenhum grupo de investidores seria capaz de consistentemente encontrar ações sub ou sobrevalorizadas.

Portanto, de acordo com Jensen (1978), não existe possibilidade de se alcançar lucros com as informações disponíveis em um mercado eficiente, principalmente pela velocidade em que o mercado é precificado. Por exemplo, no contexto do Google Trends, objeto de pesquisa e análise desse trabalho, de acordo com a HME, os dados dos índices de evolução de busca de palavras-chave no Google

se tornariam irrelevantes para a previsibilidade de movimento de mercado, uma vez que as informações úteis para a tomada de decisão dos investidores já estariam incorporadas nos preços dos ativos.

Fama (1970) categorizou a HME em três formas de eficiência, dependendo da natureza da informação disponível, sendo elas: (i) Eficiência de forma fraca: os preços dos ativos são formados pelo conjunto de informações relativas a retornos e preços do passado, nesse grau de eficiência, expectativas de preços passados não são sinalizadores eficientes de preços futuros; (ii) Eficiência de forma semiforte: os preços refletem todo o conjunto de informações públicas disponíveis a todos os participantes do mercado, por exemplo, anúncio de ganhos anuais em reportes da empresa e noticiários. Nesse grau, a informação ao se tornar pública é imediatamente incorporada nos preços dos ativos, não permitindo nenhum retorno extraordinário; (iii) Eficiência de forma forte: os preços refletem todo o conjunto de informações privadas, logo nenhum investidor conseguiria retornos anormais, mesmo utilizando dados internos e confidenciais.

Posteriormente, Fama (1991) propôs uma revisão na abordagem da eficiência de forma fraca, ao incluir metodologias de testes mais amplas de previsibilidade dos retornos, por meio da utilização de variáveis como dividendos e taxa de juros. Para as demais categorias (semiforte e forte), o autor sugeriu alterações nos títulos para estudos de eventos e testes de informações privadas, respectivamente.

O debate sobre o nível de eficiência em que os mercados operam em uma situação real é amplo e controverso. Chau e Vayanos (2008) e Maloney e Mulherin (1998) concluíram seus estudos suportando a eficiência forte dos mercados, resultados opostos aos identificados por Kyle (1985). Por mais que existam exemplos suportando esse grau de eficiência, conforme Ross et al. (1995), o consenso acadêmico é de que a forma forte é algo substancialmente teórico e possivelmente nenhum mercado é capaz de atender completamente seus requisitos, além da dificuldade em se mensurar a influência da informação no preço. Por outro lado, vários estudos suportam a forma semiforte, como Scbiehll (1996) e Malkiel (1995). Contrários a teoria do *Random Walk* e a forma fraca de eficiência, Lo e Mackinlay (1987) e Jegadeesh (1990) evidenciaram a possibilidade de se prever retornos com base no histórico do preço, ao confirmarem a correlação entre as variáveis analisadas.

A partir de 1980, pesquisadores começaram a discutir a consistência da teoria do mercado eficiente, desafiando principalmente o conceito de racionalidade completa dos investidores. O principal questionamento era se o excesso de volatilidade das ações poderia ser previsto pelo modelo de mercado eficiente, uma vez que anomalias e bolhas financeiras não tiveram explicações consistentes suportadas pela HME. Nesse contexto, Shleifer e Vishny (1997) afirmaram que, mesmo após levar em conta o risco fundamental e os custos de transação, a teoria financeira clássica não conseguiria explicar erros de avaliação prolongados e oportunidades de arbitragem ainda não exploradas.

Anomalias, na visão de Haugen (2000), são tendências de comportamento de preços que contradizem a previsão teórica aceita, no caso, a hipótese do mercado eficiente. Elton e Gruber (1995) apresentaram algumas possíveis explicações que podem ocasionar anormalidades nos mercados financeiros, como por exemplo, os mercados são simplesmente ineficientes. Bruni e Famá (1998) identificaram na literatura estudos com dados históricos que evidenciavam a existência das seguintes anomalias: efeito janeiro, efeito mudança de mês, efeito segunda-feira, relação contábil sobre valor de mercado, relação preço sobre vendas, relação lucro sobre preço, rendimento dos dividendos, ações negligenciadas, tamanho da empresa e efeito de anúncios de propaganda.

Para buscar explicações sobre a causa das anomalias e defender que o mercado financeiro é consistente com influências imperfeitamente racionais na formação de seus preços, na reação a novas notícias e na maneira como são conduzidas as negociações, economistas começam a incorporar fatores comportamentais nos modelos tradicionais, dando origem a teoria das finanças comportamentais. (HIRSHLEIFER;TEOH, 2003).

Um dos artigos iniciais dessa teoria foi publicado pelos psicólogos israelenses Kahneman e Tversky (1979). O Prospect Theory: An Analysis of Decision under Risk criticou a teoria da utilidade esperada e apresentou um modelo alternativo de tomada de decisão sob risco, chamado teoria do prospecto. Eles verificaram que as escolhas dos investidores entre os prospectos de risco exibem vários efeitos incoerentes com os princípios básicos da teoria da utilidade, incoerências motivadas pelos padrões de comportamento irracionais dos indivíduos.

O principal objetivo das Finanças Comportamentais, segundo Shiller (2003) é estudar e compreender o comportamento dos investidores, relacionando as imperfeições de mercado a uma combinação de vieses cognitivos e erros de raciocínio e processamento de informações e, a partir disso, apresentar as possíveis soluções para essa irracionalidade. De acordo com Barberis e Thaler (2002), essa abordagem é construída com a combinação de dois fatores principais: limites de arbitragem e psicologia humana. Enquanto, a psicologia aborda as consequências sistemáticas que surgem quando as pessoas tomam decisões baseadas em preferências e ao formar crenças, os limites de arbitragem estão relacionados aos limites de atuação dos agentes racionais no mercado.

De acordo com Halfeld e Torres (2001)

“O homem das Finanças Comportamentais não é totalmente racional; é um homem simplesmente normal. Essa normalidade implica um homem que age, frequentemente, de maneira irracional, que tem suas decisões influenciadas por emoções e erros cognitivos, fazendo com que ele entenda um mesmo problema de formas diferentes, dependendo da maneira como é analisado. O campo de estudos das Finanças Comportamentais é justamente a identificação de como as emoções e os erros cognitivos podem influenciar o processo de decisão de investimentos e como esses padrões de comportamento podem determinar mudanças no mercado. O grande desafio para os pesquisadores do tema está em provar que tais anomalias de comportamento são realmente previsíveis e podem modificar o mercado de forma definida”. (HALFELD;TORRES, 2001 pág. 2)

A teoria comportamental, conforme Thaler (2016) melhora a assertividade das previsões econômicas, ao considerar viesamentos e comportamentos inexplorados pela teoria das finanças moderna e, dessa forma, aproxima as análises econômicas à realidade. Dentre os padrões de comportamento cognitivos e emocionais dos investidores que afetam o funcionamento dos mercados financeiros, Fernandes (2015) cita: avaliação subjetiva do risco, aversão à ambiguidade, decisões de investimento baseadas na representatividade, conservadorismo, excesso de confiança, sentimento, otimismo excessivo ou pessimismo, atenção limitada e *herding*. O presente trabalho se concentra em aspectos relacionados a atenção e ao sentimento dos investidores.

2.2 ATENÇÃO DO INVESTIDOR

A teoria moderna das finanças assume que as informações são incorporadas instantaneamente nos preços dos ativos, visto que os investidores têm recursos cognitivos sem restrições e são sempre capazes de dar atenção adequada a todos os tipos de informações relacionadas às ações. Este pressuposto exige, segundo Da et al. (2011), que os investidores dediquem atenção suficiente aos ativos. Entretanto, Simon (1972) indicou que a capacidade cognitiva do cérebro humano para reunir e processar informações é limitada, impedindo que os indivíduos tomem decisões plenamente informados e de forma racional.

Na mesma linha, Kahneman (1973) sugeriu que: (i) a atenção é um recurso cognitivo escasso, com isso as habilidades individuais de receber e processar as informações são restringidas pelo limite da atenção; (ii) a quantidade de atenção ou esforço depende, principalmente, das demandas das atividades exercidas, sendo muitas vezes insuficiente para compensar tarefas muito complexas; (iii) a atenção é divisível, sendo sua alocação uma questão de escolha. Em altos níveis de carga de tarefas, no entanto, a atenção torna-se quase exclusiva; (iv) a atenção é seletiva, ou controlável, sendo utilizada para facilitar o processamento ou execução de percepções selecionadas.

Corroborando, Pashler (1998) argumentou que o termo atenção não se refere apenas as limitações nas percepções de estímulos, mas também a limitações gerais do funcionamento mental, tais como a tomada de decisões, o armazenamento de informações na memória e o planejamento de ações.

Partindo do pressuposto que o processamento de todas informações do mercado acionário é impraticável, o modelo de expectativas racionais de Grossman e Stiglitz (1980) sugeriu que os mercados podem não ser totalmente eficientes em termos de informação, visto que o custo de sua obtenção é alto. Os autores argumentaram que, se os preços de mercado incorporassem todas as informações disponíveis, e as informações são onerosas, então ninguém teria incentivo para adquiri-las. Por outro lado, a partir da ideia de que a informação consome atenção, uma maior atenção levaria a uma maior absorção de informações pelos investidores, tornando os mercados mais eficientes.

Em sua hipótese cognitiva do investidor, Merton (1987) pontuou que a atenção do investidor pode estar relacionada com a liquidez e preço das ações. O modelo defendeu que devido à limitação no processo de obtenção das informações, os investidores operam em uma pequena fração do número total de títulos disponíveis no mercado, ou seja, um subconjunto de ativos. Dessa forma, o investidor tende a somente comprar ações de empresas conhecidas, e não age impulsivamente na compra de ações desconhecidas. Notícias ou informações relevantes sobre a empresa ou setor, divulgadas pelos principais veículos de comunicação e amplamente distribuídas nas mídias, podem contribuir para que parte dos investidores de mercado acompanhem a empresa. Ao avaliarem as informações detalhadas sobre a empresa, se tornariam novos acionistas, não de maneira impulsiva, e, por consequência, o valor da empresa aumentaria.

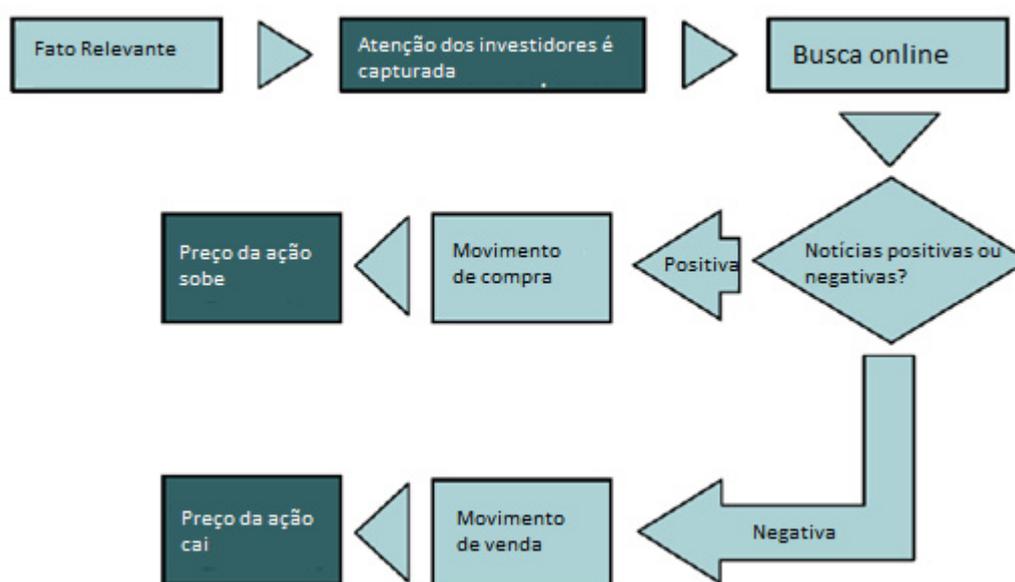
Para reparar a assimetria de informações, o autor sugeriu que ações de empresas com uma base pequena de investidores, portanto menos conhecidas, tendem a ter retornos superiores as ações de empresas mais conhecidas. Na lógica do estudo, seguindo uma estratégia de investimento inclinada para esses ativos, os investidores poderiam obter retornos acima da média de mercado. A medida que a base de investidores é aumentada, podendo o período de tempo ser considerável ou não, esse retorno em excesso tende a ser eliminado. Dessa forma, Merton (1987) concluiu ressaltando que os custos de informação podem causar comportamentos anômalos em um mercado dominado por agentes racionais, porém, seja no curto ou no longo prazo, essa ineficiência tende a ser eliminada ou minimizada.

Odean (1999) desenvolveu as ideias de Merton (1987) e sugeriu que os investidores concentram o problema de escolher um título dentre vários disponíveis, limitando suas pesquisas a ações que chamaram a atenção recentemente. Nessa linha, mais recentemente, ao criticarem o trabalho de Grossman e Stiglitz (1980), relatando que os autores utilizaram a premissa de que a decisão de compra e venda dos investidores seguem os mesmos pressupostos (divergindo apenas por um sinal de mais ou menos nas equações do modelo), Barber e Odean (2008) introduziram a hipótese "*price pressure hypothesis*", ao verificarem que investidores individuais são mais propensos a comprar ações que captam suas atenções, causando uma pressão positiva temporária sobre os preços, possivelmente revertida quando a pressão acabar por se dissipar. A impossibilidade de se analisar todas as ações disponíveis

para compra, leva o investidor a escolher aquelas que chamam sua atenção, comportamento não verificado no processo de venda, em razão da opção limitada de ações do portfólio disponíveis para venda. Além disso, os autores identificaram que o comportamento de compra de investidores individuais é mais influenciado pela atenção do que o comportamento de compra de investidores profissionais.

Alternativamente a teoria de Barber e Odean (2008), Tantaopas et al. (2016) com base nos resultados de Vozlyublennaia (2014) e Vlastakis e Markellos (2012), defendem que a pressão temporária de preços ocorre de maneira positiva ou negativa, uma vez que a descoberta de informações, causadas pelo aumento da atenção dos investidores, influenciam tanto o lado comprador quanto o lado vendedor do mercado, conforme fluxo apresentado na Figura 1.

FIGURA 1 - Diagrama do efeito da atenção na incorporação das informações nos preços dos ativos.



Fonte: Guzella (2020) com base em Tantaopas et al. (2016)

No campo da liquidez de mercado, Corwin e Coughenour (2008) concluíram que devido a atenção limitada, os investidores dedicam mais esforços às suas ações mais ativas, diminuindo a liquidez nos títulos com menor atenção associada.

Em uma literatura crescente, autores enfatizam o comportamento dos agentes de mercado, partindo da premissa de que a capacidade limitada de atenção dos indivíduos pode impedi-los de tomar decisões totalmente racionais, levando a resultados não otimizados. Por exemplo, Nieuwerburgh e Veldkamp (2010) sugeriram

que quando os investidores têm que escolher quais informações devem coletar antes de investir, a alocação de sua carteira não é otimizada em termos de diversificação. Peng et al. (2005) verificaram que devido à escassez de atenção, investidores tendem a processar mais informações do mercado (exemplo taxa de desemprego, inflação, entre outros) do que informações específicas da empresa, levando a uma análise mais genérica das possibilidades de investimento. Nesse cenário, a falta de informações gera nos investidores um excesso de confiança, tornando suas decisões irracionais ao negligenciar detalhes importantes ou ao realizar análises financeiras de qualidade não satisfatória.

A partir da ideia que investidores se concentram apenas nas informações de seus interesses imediatos, Huberman e Regev (2001) estabeleceram que novas informações somente afetam o preço das ações quando os investidores de fato prestaram atenção a elas. Os autores verificaram que as ações da empresa EntreMed quadruplicaram após a publicação de um artigo no jornal The New York Times sobre uma possível cura para o câncer, entretanto, cinco meses antes, essa informação já havia sido noticiada na revista Nature. Caso similar foi apresentado por Engelberg et al. (2012), ao verificarem que recomendações de ações anunciadas no programa de televisão Mad Money geram um choque de atenção nos telespectadores, ocasionando um acréscimo nos preços das ações, que se revertem meses subsequentes.

Isto posto que a atenção limitada afeta o desempenho do mercado financeiro, quantificá-la pode ser um processo desafiador aos pesquisadores e analistas, por se tratar de um fenômeno abstrato e não quantificável. Na literatura, *proxies indiretas* são utilizadas para tentar medir a atenção dos investidores. No estudo de Barber e Odean (2008), por exemplo, a atenção dos investidores foi quantificada por retornos extremos, volumes anormais de negociação e a presença ou não das empresas analisadas em noticiários ou manchetes, identificando que os investidores são compradores líquidos em dias de alto volume e são influenciados por retornos extremamente negativos/positivos de dias anteriores e por quando as ações estão nos noticiários. Grullon et al. (2004) verificaram que despesas com propaganda afetam o retorno das ações, a liquidez do ativo e a base de acionistas, resultados similares aos identificados por Chemmanur e Yan (2009). Fang e Peress (2009) argumentou que ações sem a cobertura da mídia ganham maior retorno do que ações com alta mídia. Seasholes e Wu (2007) concluíram que as ações que atingem os limites superiores

de preço têm retornos significativamente mais altos, porém revertidos nas semanas posteriores. Por fim, utilizando os dias da semana como *proxies*, Dellavigna e Pollet (2005) identificaram que, nas sextas feiras, os investidores têm uma resposta fraca as novas informações de mercado.

As *proxies* indiretas, entretanto, não asseguram a atenção do investidor em determinados eventos. Por exemplo, de acordo com Da et al. (2011), manchetes de jornais não confirmam que os agentes de mercado a observaram, a menos que eles a leiam efetivamente, bem como o volume de negociações e retornos extremos podem ser causados por fatores não relacionados com a atenção dos investidores.

Com o objetivo de superar estes desafios e fornecer evidências mais diretas ao efeito de atenção limitada, a literatura recente sugere a utilização de *proxies* diretas, capazes de assegurar que os investidores estão prestando atenção a algum fato. Por exemplo, são elaboradas com base em entrevistas com investidores, preenchimento de questionários ou análise de pesquisas em sites de busca (exemplo: Google, Yahoo!, AOL).

2.3 SENTIMENTO DO INVESTIDOR

Assim como na abordagem da atenção limitada dos investidores, a teoria moderna das finanças, ao considerar o indivíduo racional, elimina os efeitos causados pelo sentimento dos investidores. Baker e Wurgler (2007), por exemplo, sugeriram que nessa teoria, os investidores não-emocionais sempre forçam os preços do mercado de capitais a serem equivalentes ao valor presente racional dos fluxos de caixa futuros esperados.

Zhang (2008) afirmou que a definição de sentimento do investidor é tratada de várias formas na literatura, enquanto alguns pesquisadores se referem ao como uma propensão a negociar com ruído em vez de informações, outros utilizam o termo para se referir ao otimismo e pessimismo do investidor ou sua aversão ao risco. Zweig (1973) argumentou que o sentimento do investidor vem das expectativas tendenciosas dos investidores sobre o valor dos ativos. Segundo Baker e Wurgler (2007), é uma percepção sobre fluxos de caixa futuros e riscos de investimentos e não justificada

pelos seus fundamentos. Para Asif Khan et al. (2019), sentimento é a forma que os investidores desenvolvem suas preferências e crenças utilizando humores, emoções, preconceitos psicológicos e cognitivos.

Keynes (1936) já havia destacado o papel da psicologia e do sentimento na economia, ao sugerir que as flutuações de mercado podem ser influenciadas pelo espírito animal dos investidores (*animal spirits*), construído quando indivíduos com sentimentos altos (baixos) tendem a fazer julgamentos e escolhas excessivamente otimistas (pessimistas). A partir do pensamento de Keynes, a literatura começou a formalizar o papel do sentimento do investidor no mercado financeiro.

Um dos trabalhos pioneiros é o modelo de “*noise traders*” desenvolvido por De Long et al. (1990). No modelo proposto, os investidores são classificados em duas categorias: investidores racionais (arbitradores que atuam com base em informações disponíveis) e os investidores irracionais ou *noise traders* (indivíduos que atuam baseados em sentimentos ou ruídos de mercado), esses dois perfis competem nos mercados financeiros e estabelecem preços e expectativas de retorno. Uma das principais conclusões do modelo é que o risco sistemático, não diversificável, da imprevisibilidade das ações dos investidores irracionais limita a atuação dos arbitradores do mercado. Como consequência, os preços de mercado dos títulos se afastam consideravelmente de seus valores fundamentais e geram uma volatilidade excessiva. A atuação de investidores irracionais e os limites de arbitragem podem explicar anomalias nos mercados financeiros, como, por exemplo, bolhas especulativas. Complementando, Shleifer e Vishny (2005) sugeriram que, em algumas circunstâncias, os arbitradores podem não ter capacidade para corrigir os preços para os seus valores fundamentais, tais como limitações de capital ou custos de agência.

Brown e Cliff (2005) descreveram que os sentimentos possuem papel relevante na formação dos mercados financeiros, por dois motivos principais: (i) o sentimento pode ser considerado como uma variável persistente ao longo do tempo, pois a medida que as pessoas se tornam otimistas (pessimistas), cada vez mais pessoas tendem a aderir a tendência; (ii) é provável que a arbitragem elimine os preços incorretos a curto prazo, porém é incapaz de corrigir incompatibilidades em horizontes mais longos. Corroborando, Fischer Black (1986) afirmou que os investidores negociam títulos com base nos ruídos de mercado ao invés das informações disponíveis.

Ao interpretar os estudos de Baker e Wurgler (2007), Brochado (2019) apontou três conclusões predominantes nos estudos do sentimento dos investidores: (i) o retorno das ações pode ser explicado pelo sentimento dos investidores, por meio de uma relação negativa entre sentimento e retorno. (ii) o sentimento tem maior influência sobre as ações cujas valorizações são mais subjetivas e difíceis de serem objeto de arbitragem (por exemplo, ações com maior volatilidade e maior risco de queda); (iii) a influência do sentimento do investidor está sujeita a reversões. Corroborando, Schmeling (2009) e Fisher e Statman (2000) verificaram que quando o sentimento dos investidores é alto, os retornos das ações tendem a ser menores, e vice-versa.

De acordo com Baker e Wurgler (2007), a questão central não é mais se o sentimento do investidor afeta o preço das ações, mas como medir o sentimento do investidor e quantificar seus efeitos. Assim como na medição da atenção dos investidores, captar o sentimento dos investidores é uma tarefa difícil, visto a sua irracionalidade. Pesquisadores têm apresentado muitas alternativas para calcular e quantificar o sentimento do investidor. Tradicionalmente, conforme Da et al. (2015), o sentimento dos investidores é medido de duas formas: (i) baseada em informações do mercado (forma indireta) e (ii) baseado em pesquisas de opinião/comportamento de investidores (forma direta). A forma indireta utiliza principalmente dados financeiros e econômicos disponíveis, como: liquidez, prêmio de dividendos, quantidade de IPOs e seus retornos no primeiro dia de negociação, volume de negociações, descontos de fundos fechados (Lee et al., 1991) e volatilidade (Whaley, 2000). Apesar dos dados utilizados como *proxies* estarem disponíveis em uma frequência relativamente alta, essa métrica pode ser enviesada por outros fatores além do sentimento dos investidores.

Nesse contexto, objetivando medir o sentimento do investidor de uma forma mais conclusiva e com menos ruídos, Baker e Wurgler (2006) ao agruparem seis indicadores mercado, propuseram a criação de um índice indireto de sentimento de mercado. Os indicadores utilizados foram: taxa de descontos de fundos fechados, rotatividade de transações, número de empresas com IPO, retorno no primeiro dia de negociação, número de novas ações emitidas e prêmio de dividendo.

A abordagem que utiliza a opinião/comportamento dos investidores é considerada como uma medida explícita do sentimento, porém vulnerável e limitada a

fatores como: qualidade da resposta do entrevistado (muitas vezes possui pouco incentivo para responder adequadamente), frequência de obtenção de dados baixa e alto custo de implementação. As pesquisas de sentimento (Brown e Cliff, 2005; Fisher e Statman, 2000) e índices de confiança dos consumidores (Lemmon e Portniaguina, 2002; Schmeling, 2009; Fisher e Statman, 2000) são exemplos de *proxies* utilizados para medir o sentimento de forma direta.

A análise de dados textuais das mídias sociais e notícias é outro método que vêm ganhando destaque na literatura. Por exemplo, Tetlock (2007) mediu o sentimento dos investidores pela frequência das palavras positivas e negativas em uma coluna diária do jornal Wall Street Journal's. Bollen et al. (2011), Siganos et al. (2017) e Moat et al. (2013) mediram o sentimento do público por meio de dados do Twitter, Facebook e do Wikipédia, respectivamente.

Recentemente, métodos para medir o sentimento dos investidores usando pesquisas na Internet foram adotadas. Renault (2017) afirma que a extração de dados da rede pode ser eficaz, uma vez que elimina questões relacionadas aos viesamentos em índices baseados em pesquisas, componentes idiossincráticos não relacionados com o sentimento em indicadores de mercado e causalidade confusa gerada por notícias. A principal referência nessa abordagem é o artigo publicado por Da et al. (2015) e detalhado no tópico a seguir.

2.4 GOOGLE TRENDS E A ATENÇÃO E SENTIMENTO DOS INVESTIDORES

O entendimento sobre os potenciais impactos da atenção e sentimento dos investidores nos mercados foram aprofundados a medida em que ferramentas de análise de grandes quantidades de dados foram disponibilizadas ao público, como por exemplo, o Google Trends (GT).

O Google Trends é uma ferramenta gratuita disponível ao público, de propriedade da empresa Alphabet Inc, que fornece o quão frequente palavras chaves estão sendo procuradas pelos usuários do Google, permitindo filtrar os resultados de acordo com a localização, categoria da palavra e o intervalo de tempo. É possível gerar dados em tempo real que facilitam a identificação e a compreensão de

tendências de comportamento, do sentimento da população sobre determinados temas e o entendimento de fenômenos pontuais.

O GT tem sido utilizado frequentemente em pesquisas na área da saúde pública, da política, do turismo, da economia, das finanças e entre outros temas. Por exemplo, Choi e Varian (2011) encontraram, por meio de modelos autorregressivos, relação positiva entre o volume de buscas e as vendas de automóveis, residências e destinos turísticos, no mercado norte americano, entre 2006 e 2011. Conforme, Carneiro e Mylonakis (2009) com dados americanos entre 2004 e 2009, foi possível rastrear o agravamento da doença Influenza, pela análise de buscas relacionadas aos principais sintomas associados a enfermidade. Schmidt e Vosen (2009) demonstraram que os dados fornecidos pelo GT podem servir de parâmetro para o cálculo de índices de confiança dos consumidores.

A abordagem baseada em pesquisas de termos na internet começou a ganhar atenção na literatura financeira após os trabalhos publicados por Da et al. (2011) e Mondria et al. (2010).

Da et al. (2011) sugerem uma nova forma de medir a atenção dos investidores do mercado financeiro, por meio da utilização do índice de evolução de buscas no Google (GSVI, do inglês *Google Search Volume Index*), escolha justificada pelos seguintes motivos: o Google é a ferramenta de busca mais utilizada nos Estados Unidos; quando uma ação é procurada no site, é inquestionável que se está prestando atenção nisso; e a base de pesquisas possui potencial para fornecer dados *real time* das atividades econômicas.

Utilizando como parâmetro de pesquisa os *tickers* (códigos de negociação dos títulos no mercado) de papéis listados no índice norte americano Russell 3000, entre 2004 e 2008, os autores desenvolveram o estudo em três principais etapas, sendo elas: (i) análise da correlação e da relação temporal, via modelos de vetores autorregressivos (VAR), entre as GSVI e as medidas de atenção até então predominantes na literatura, como: retornos anormais das ações, volume de negociação, número de notícias divulgadas e gastos com publicidade e propaganda, (ii) análise de ordens de negociação extraídas da SEC, para verificação se de fato as GSVI podem capturar a atenção de investidores individuais (iii) teste da hipótese de Barber e Odean (2008), a qual sugere que investidores individuais negociam as ações

que chamam suas atenções e causam a pressão positiva nos preços, através de regressões *cross-section*.

Dentre as conclusões apresentadas por Da et al. (2011), destacamos que a o índice de evolução de pesquisas no Google possibilita a captura da atenção de investidores, de forma mais eficaz que as medições até então utilizadas, sugerindo que um aumento no volume de buscas prediz um incremento de preço das ações nas duas semanas seguintes e uma eventual reversão dentro de um ano, confirmando a hipótese de Barber e Odean (2008), especialmente entre as ações de pequeno porte e as negociadas por pequenos investidores e, dessa forma, demonstrando que a ferramenta possui potencial preditivo de comportamento de mercado, desafiando as ideias fundamentais da teoria do mercado eficiente de Eugene Fama.

Ao ser aceita como uma medida relevante para a atenção e sentimento dos investidores, vários outros estudos começaram a analisar a GSVI para compreender o comportamento dos mercados financeiros, principalmente no que tange indicadores financeiros de volume negociado, variação nos preços e volatilidade das ações.

Com base nas contribuições de Da et al. (2011), Joseph et al. (2011) utilizou as intensidades de buscas dos *tickers* de empresas no Google, o modelo de três fatores de Fama e French (1993), sendo eles: (a) retorno em excesso do título em relação ao mercado; (b) diferença entre os retornos de um portfólio diversificado com ações de baixo valor de Mercado e alto valor de mercado e; (c) diferença de retorno entre uma carteira com os indicador “Book to Market” (Preço da ação no mercado/Valor contábil por ação) alto e uma carteira com o indicador baixo, e as variáveis de momento sugeridas por Carhart (1997) para entender as relações entre as *proxies* e os retornos anormais e excessivos volumes de negociação das empresas listadas na S&P500, entre 2005 e 2008. A metodologia adotada para obter resultados mais robustos, foi de classificar semanalmente o grupo de empresas em cinco quintis, conforme evolução da intensidade de busca em relação à semana anterior e, para cada um desses quintis, calcular, por exemplo, a regressão do modelo de quatro fatores (Fama e French + Carhart) é representada pelo modelo econométrico 2.1.

$$R_{pt} - R_{ft} = \alpha + \beta_m(R_{mt} - R_{ft}) + \beta_sSMB_t + \beta_hHML_t + \beta_uUMD_t + \varepsilon_t \quad (2.1)$$

Onde,

R_{pt} : retorno do ativo ou portfólio p no período t

R_{ft} : é o retorno do ativo livre de risco no período

$R_{mt} - R_{ft}$: o retorno em excesso em relação ao mercado.

SMB_t : diferença entre os retornos de um portfólio diversificado com ações de baixo valor de Mercado e alto valor de mercado

HML_t : diferença de retorno entre uma carteira com o indicador “Book to Market” alto e baixo.

UMD_t : a diferença de retorno entre uma carteira de ações com alto retorno no ano passado e uma carteira de ações com baixo retorno no ano passado.

Após compararem os resultados para cada um dos quartis, Joseph et al. (2011) concluíram que maiores índices relativos de evolução de buscas estão associadas a maiores retornos nas semanas subsequentes, *tickers* com maior intensidade GSVI possuem um volume anormal de negociação 158% superior às empresas com menor intensidade de pesquisa, além do efeito preditivo da volatilidade ser mais visível em ações difíceis de arbitrar e de maiores riscos.

Similarmente, Mondria e Wu (2012) ao analisarem as ações do mesmo índice, porém com uma frequência mensal e dados entre 2004 e 2009, identificaram que os títulos associados a um aumento de atenção obtêm retornos mais elevados no mês seguinte. Entretanto, os resultados sobre a relação entre o volume de pesquisa do Google e os retornos são heterogêneos na literatura, Bijl et al. (2016), por exemplo, sugerem, com dados mais recentes entre 2008 e 2013, indicativos opostos aos apresentados por Da et al. (2011), Joseph et al. (2011) e Mondria e Wu (2012), ao concluírem que aumentos nos níveis de pesquisas são sucedidos de quedas nos retornos das ações listadas na S&P500. Preis et al. (2010), por outro lado, não identificaram correlações entre as variações de buscas e os movimentos nos preços de compra e venda de ações do índice S&P500, verificando apenas uma relação forte com volumes transacionados.

Enquanto os resultados empíricos entre atenção do investidor e retornos são contraditórios, as análises entre atenção e volume de negociação e volatilidade indicam uma forte relação entre as variáveis.

Com base em dados fornecidos pelo Yahoo!, Bordino et al. (2012) analisaram a relação entre o volume de negociação das principais ações listadas na NASDAQ e as pesquisas dos respectivos códigos (*tickers*) de negociação. Após análises de correlação de Pearson e causalidade de Granger, os resultados apontam que um aumento no tráfego de pesquisas dos *tickers* sugere um aumento nos volumes negociados das ações no dia seguinte, corroborando com a tese de que é possível medir a atenção dos investidores por meio do seu comportamento em sites de buscas.

Vlastakis e Markellos (2012) analisaram as 30 empresas mais negociadas nas bolsas NASDAQ e NYSE, objetivando entender a influência dos volumes de pesquisas (*information demand*) e de notícias divulgadas na mídia (*information supply*) nas variáveis volatilidade e volume de negociação das firmas e do mercado. Os resultados empíricos indicam que ambas as *proxies* estão significativamente associadas, além de confirmarem a hipótese de que as procuras por informações das empresas e dos mercados aumentam em períodos de grande volatilidade e elevados retornos. Resultados similares sobre a volatilidade foram identificados por Dimpfl e Jank (2011) ao analisarem os índices Dow Jones (DJIA), FTSE 100, CAC 40 e DAX.

Complementando os estudos de Vlastakis e Markellos (2012), Andrei e Hasler (2013) incluíram, junto à atenção dos investidores, os impactos das incertezas de mercado, também concluindo que as variações do retorno e dos riscos das ações estão relacionadas com variações na atenção e na incerteza.

Kristoufek (2013) propôs a criação de uma estratégia de diversificação de ações nos portfólios de investimento e concluiu que as informações do Google Trends podem servir de ferramenta de auxílio na minimização da exposição ao risco, confirmando a hipótese de que os volumes de buscas no Google estão correlacionados com os graus de riscos das ações.

Vozlyublennaia (2014) argumentou que é pouco provável que investidores individuais pesquisem informações sobre empresas específicas, uma vez que o conjunto de opções oferecidas pelos intermediários sugerem investimentos em índices de mercado ou carteiras já selecionadas, além dos investidores profissionais possivelmente não procurarem as empresas no Google, visto que as plataformas de

trading oferecem a opção de leitura dinâmica de notícias. Isto posto, diferentemente dos artigos que avaliavam cada ação individualmente, o estudo sugere analisar os retornos e volatilidade para índices de mercado (ouro, petróleo, títulos públicos, S&P500, NASDAQ E DOW). A análise considerou dados semanais entre 2004 e 2012.

Primeiramente, Vozlyublennaia (2014) executou testes de causalidade de Granger entre o retorno (volatilidade) dos índices e os volumes de busca com defasamentos e, na sequência, executou regressões VAR com a adição de variáveis macroeconômicas, como, por exemplo, títulos americanos de 1 ano e rendimentos de dividendos agregado e variáveis que medem a interação defasada entre atenção e retorno.

Dentre os principais resultados identificados, destacamos (i) o aumento da atenção dos investidores afeta significativamente o retorno em um curto intervalo de tempo, podendo esse impacto ser positivo ou negativo, dependendo da natureza das informações. Essa observação diverge da hipótese da pressão positiva dos preços desenvolvida por Barber e Odean (2008); (ii) por meio da análise inversa da relação entre atenção e retorno, declínios nos resultados de mercado ou choques de retorno geram um aumento na atenção, sendo seu efeito suportado por um longo intervalo de tempo.

Como a atenção é influenciada pelos retornos, o retorno passado ou aumento (declínio) da atenção pode indicar a natureza das informações recebidas pelos investidores, o que, por sua vez, influencia o retorno futuro no curto prazo. Ao analisar esse mecanismo, Vozlyublennaia (2014) concluiu que choques de atenção reduzem a previsibilidade do retorno, aumentando a eficiência de mercado.

Grande parte da literatura concentra-se principalmente em analisar relações com o mercado financeiro norte-americano, embora estudos em outros mercados relevantes tenham sido publicados. Para as ações alemãs negociadas na plataforma Xetra, Bank et al. (2011) concluíram que um maior GSVI conduz a uma maior atividade de negociação, maior liquidez das ações e maiores retornos no curto prazo, resultados similares aos identificados por Chen (2011) e Aouadi et al. (2013) nos mercados holandês e francês respectivamente. Para o índice europeu Stoxx 50, Latoeiro (2012) identificou resultados similares aos observados por Bijl et al. (2016), ao notar que um maior índice de evolução de busca está associado a menores rentabilidades. Takeda

e Wakao (2014) identificaram que a relação da GSVI com os retornos anormais do mercado japonês é bem menos intensa se comparada com o impacto na liquidez.

Proposto por Preis et al. (2013), um novo método utilizou termos financeiros relacionados as preocupações do consumo familiar para tentar entender o sentimento dos mercados financeiros, complementando as análises anteriores que utilizavam como palavras chaves apenas os nomes das empresas, os códigos de negociação e os nomes dos índices de mercado. Os autores investigaram a relação do volume de buscas de 98 palavras no Google, como “*hedge*”, “*debt*”, “*dividends*”, “*inflation*”, “*bonds*”, “*earnings*” e “*market*”, com o desempenho do mercado acionário norte americano, no período entre 2004 e 2011. Os resultados sugerem que a variação semanal nos termos de busca fornece *insights* sobre as tendências de comportamento dos agentes econômicos, podendo ser explorados na elaboração de portfólios de investimentos rentáveis, inclusive com retornos acima dos principais *benchmarks* de mercado.

Na mesma linha de raciocínio, Perlin et al. (2016) selecionaram as 15 palavras mais utilizadas em livros referências no tema de finanças e analisaram, com dados entre 2005 e 2014, a relação de suas pesquisas semanais com a volatilidade, o retorno e o volume negociado em 4 mercados de capitais falantes da língua inglesa (Canadá, Estados Unidos, Reino Unido e Austrália). Por meio da utilização do modelo de auto regressão estrutural vetorial (VAR) foi identificado que, por exemplo, um aumento nas buscas das palavras em inglês “*finance*”, “*stock*” e “*debt*” pode impactar negativamente o retorno de índices de mercado no futuro. Adicionalmente, os autores verificaram que estratégias de investimentos utilizando a evolução do índice de busca da palavra “*stock*” como indicador de compra ou venda de ações, podem resultar em rentabilidades superiores as estratégias baseadas no *buy and hold*. Dessa forma, corroborando com os resultados identificados por Preis et al. (2013), os autores concluíram que é possível utilizar o Google Trends para auxiliar a previsão do comportamento do mercado acionário.

Ainda nesse contexto, questionando o modelo sugerido por Preis et al. (2013), Challet e Ayed (2013) identificaram, ao incluir vieses como custos de transação, que palavras financeiras não possuem poder preditivo de retorno de ações superior ao de palavras randômicas relacionadas a doenças, carros e jogos eletrônicos.

Em uma abordagem recente, Da et al. (2015) elaboraram um índice diário de sentimento dos investidores, nomeado de índice FEARS (*Financial and Economic Attitudes Revealed by Search*), desenvolvido a partir da análise da evolução da frequência de buscas, no Google, de termos financeiros relacionados as preocupações econômicas das famílias americanas. Os autores defendem que as pesquisas na web têm potencial de revelar sentimentos de maneira mais eficiente que os índices de confiança dos investidores e de que os próprios indicadores de desempenho de mercado.

A metodologia proposta por Da et al. (2015) para a construção do índice FEARS segue os seguintes procedimentos:

- 1- Identificação dos termos financeiros capazes de revelar o sentimento das famílias americanas em relação à economia, como, por exemplo, crise, ouro, inflação, recessão e falência.
- 2- Ao inserir cada palavra no Google Trends, a ferramenta retorna uma lista com mais 10 palavras que possuem associação positiva com a palavra-chave. As 149 palavras selecionadas inicialmente no estudo, geraram 1.490 termos, que, após a remoção de registros duplicados, totalizaram 1.245 termos.
- 3- A remoção dos termos não relacionados a finanças ou com base de dados insuficiente para análise resultaram em uma listagem de 118 palavras.
- 4- Para cada uma das palavras, a evolução da frequência de buscas entre 2004 e 2011 foi obtida pelos autores, incluindo o cálculo de sua variação ao longo do tempo e a eliminação dos efeitos da sazonalidade.
- 5- O índice FEARS é consolidado após o agrupamento dinâmico das 30 palavras que, após regressões retroativas *backward rolling*, obtiveram os menores *t-statistic* na relação entre a evolução diária da frequência de buscas no Google e o retorno diário de mercado. Dessa forma, a estrutura do índice pode ser representada conforme equação 2.2:

$$FEARS_t = \sum_{i=1}^{30} R^i(\Delta ASV I_t) \quad (2.2)$$

Onde,

$FEARS_t$ = Índice de Sentimento na semana t .

$R^i (\Delta ASVI_t)$ = variação no índice semanal de buscas da palavra-chave que teve um ranqueamento da classificação t-estatística, variando do menor t-statistic ($i = 1$) para o maior ($i = 118$).

Utilizando dados do índice S&P500, entre 2004 e 2011 e atualizando a composição do agrupamento das palavras do índice de sentimento a cada 6 meses, os autores analisaram econometricamente a relação entre o retorno e a volatilidade do mercado com o índice FEARS e as variáveis de controle (CBOE índice de volatilidade, EPU índice de incertezas políticas e ADS índice de condições macroeconômicas). Os resultados indicaram que o índice de sentimento pode auxiliar na previsão de retornos no curto prazo e de aumentos na volatilidade do mercado. Os resultados estão alinhados com a hipótese desenvolvida por De Long et al. (1990), principalmente ao verificar erros de preços temporários induzidos pelo sentimento.

Outros artigos replicaram a metodologia do índice FEARS, como Kostopoulos et al. (2020) para o mercado alemão, Asif Khan et al. (2019) para o mercado islâmico e Streicher (2019) para o mercado português, obtendo resultados similares.

2.5 GOOGLE TRENDS E O MERCADO FINANCEIRO BRASILEIRO

No âmbito da relação entre o mercado financeiro brasileiro e o uso dos dados de buscas na internet, a literatura ainda é escassa. Os levantamentos realizados indicaram os trabalhos de Ramos et al. (2017), Rodolfo et al. (2017) Miragaya (2017), Azevedo (2017), Viana (2017) e Caraça (2019).

Ramos et al. (2017) analisaram com qual intensidade a frequência de buscas de índices de mercado e códigos de negociação pode explicar as variáveis retorno financeiro, o volume transacionado e a volatilidade do índice Ibovespa e das ações PETR4 e VALE5, para o período entre 2007 e 2014. Estimando testes de causalidade de Granger, baseados em modelos econométricos VAR e com a utilização dos dados

do Google Trends dessazonalizados, os autores identificaram a existência de um efeito preditivo entre a frequência de pesquisas e as variáveis financeiras, destacando: (i) um aumento da atenção no *ticker* PETR4 precede um maior volume de transações, (ii) maiores frequências de buscas do índice de mercado está associada a retornos positivos e (iii) quando os retornos dos títulos aumentam, os investidores buscam mais informações sobre estes ativos. Com metodologia semelhante, Caraça (2019) analisou os dados do Ibovespa e da empresa Magazine Luiza (código de negociação MGLU3) entre 2014 e 2018, obtendo resultados similares aos encontrados por Ramos et al. (2017).

As hipóteses de que as negociações no mercado brasileiro são impulsionadas por mudanças na atenção dos investidores e que as pressões temporárias sobre os preços dos bens são afetadas por uma atenção anormal dos investidores foram testadas por Rodolfo et al. (2017). Por meio de regressões com dados em painéis de empresas listadas na B3, período entre 2010 e 2015, os autores confirmaram que a frequência de buscas no Google explica significativamente e positivamente a variação no volume de negociação, porém não justificam uma pressão temporária nos preços.

Viana (2017) examinou a correlação semanal entre as pesquisas de 120 empresas listadas na B3 e quatro variáveis de mercado (quantidade de negociações, quantidade de títulos, volume de negociação e preço da ação), para o período entre 2012 e 2017. Dentre as correlações calculadas, 57% apresentaram um p-valor (correlação) significativo, sendo a variável preço a de maior correlação com os volumes de busca. Defasando os dados extraídos do Google Trends em uma semana, as correlações significativas aumentam para 63%.

Analisando a relação reversa entre a atenção dos investidores e o mercado financeiro, Miragaya (2017) confirmou, por meio de regressões lineares entre o logaritmo das buscas no Google e retornos anormais, a hipótese de que retornos anormais das ações listadas no Ibovespa levam investidores a buscarem informações no Google, sendo a influência dos retornos negativos superior aos efeitos dos retornos positivos. Os resultados indicaram uma persistência de até 3 semanas nos efeitos dos retornos sobre a atenção dos investidores. Para a execução do trabalho foram coletados dados semanais, entre 2007 e 2016, de 58 empresas que compunham o índice Ibovespa em dezembro de 2016

Por fim, com base no trabalho de Da et al. (2015), Azevedo (2017) identificou que buscas de palavras com conotações positivas e negativas e relacionadas com economia podem fornecer informações relevantes sobre o impacto do sentimento das pessoas na economia norte americana, em especial, no mercado de ações, na curva de desemprego e na produção industrial. Entretanto, considerando o escopo brasileiro, os resultados não indicaram relações estatisticamente significantes do conjunto de palavras em português com as variáveis macroeconômicas do país. O autor levanta a hipótese de que os resultados podem ter sido afetados pela tradução simples das palavras extraídas em inglês do trabalho de Da et al. (2015) e do dicionário Harvard IV-4.

3 METODOLOGIA

Nesta seção, descrevemos o escopo do trabalho, a metodologia de definição da amostra de dados utilizada, as principais métricas selecionadas e as análises empíricas sugeridas para desenvolvimento da pesquisa.

3.1 ESCOPO DO TRABALHO E DADOS

O escopo de análise do estudo compreende o período entre janeiro de 2015 e dezembro de 2019, totalizando 5 anos, e tem como abrangência o mercado de ações do Brasil. Entende-se que esse intervalo corresponde uma janela relevante do desempenho econômico brasileiro, incluindo momentos de aceleração da economia e períodos de crises financeiras.

Para analisar as hipóteses e os objetivos propostos, a seleção dos dados foi baseada em 3 principais frentes de pesquisa, sendo elas: principais ações listadas na bolsa de valores B3, índices de ações do mercado brasileiro e palavras relacionadas aos sentimentos e comportamento dos investidores e das famílias brasileiras. A seguir são detalhados os escopos de cada um desses pilares de análise.

3.1.1 Ações listadas na B3

Inicialmente, selecionamos os 100 códigos de ações que compunham o índice IBrX100 (Índice Brasil) com a data base outubro de 2020. O índice mede o desempenho médio das cotações dos 100 ativos de maior negociabilidade e representatividade do mercado de ações brasileiro. A decisão de analisar as ações a partir desse índice, teve como objetivo assegurar que as movimentações mais relevantes do mercado fossem refletidas nas análises empíricas realizadas. Conforme informações disponíveis no site da B3, em outubro de 2020, o índice representava cerca de 88% do valor de mercado das ações negociadas, totalizando R\$ 3,55 trilhões de valor de mercado. Apesar dos participantes do índice serem ajustados

periodicamente ao longo do período analisado, optou-se por escolher a data de outubro de 2020 como data de corte do índice em razão da disponibilidade dos dados. É importante salientar como limitação do trabalho, que essa escolha de data pode enviesar a análise, visto que apenas ações “vencedoras” e relevantes tendem a se manter no índice.

Excluimos da amostra 26 títulos de ações, em razão da impossibilidade de se coletar dados para todo o período do escopo. Em 21 casos, a abertura de capital (IPO) foi realizada dentro do período entre 2016 e 2020 e, em 5 casos, o código de negociação foi alterado ao longo do período analisado ou os dados de retorno estavam incompletos.

Isto posto, a amostra de ações listadas foi composta por 74 títulos, representando 83% de participação no índice IBrX100. O Quadro 1 detalha os códigos de negociação, os nomes das empresas, o tipo de ação e os setores e segmentos de atuação.

QUADRO 1 – Empresas selecionadas na amostra via índice IBRX-100

| Código | Empresa | Tipo | Setor e Segmento |
|---------------|-----------------|-------------|--|
| ITUB4 | ITAUUNIBANCO | PN ED N1 | Financeiro e Outros |
| PETR4 | PETROBRAS | PN N2 | Petróleo, Gás e Biocombustíveis |
| ITSA4 | ITAUSA | PN ED N1 | Financeiro e Outros |
| ABEV3 | AMBEV S/A | ON | Consumo Não Cíclico / Bebidas |
| BBDC4 | BRADESCO | PN EJ N1 | Financeiro e Outros |
| PETR3 | PETROBRAS | ON N2 | Petróleo, Gás e Biocombustíveis |
| VALE3 | VALE | ON NM | Materiais Básicos / Mineração |
| B3SA3 | B3 | ON NM | Financeiro e Outros |
| VVAR3 | VIAVAREJO | ON NM | Consumo Cíclico / Comércio |
| JBSS3 | JBS | ON NM | Consumo Não Básico / Alimentos Processados |
| BBAS3 | BANCO DO BRASIL | ON NM | Financeiro e Outros |
| BBDC3 | BRADESCO | ON EJ N1 | Financeiro e Outros |
| CCRO3 | CCR SA | ON NM | Bens Industriais / Transporte |
| CIEL3 | CELO | ON NM | Financeiro e Outros |
| UGPA3 | ULTRAPAR | ON NM | Petróleo, Gás e Biocombustíveis |
| EQTL3 | EQUATORIAL | ON NM | Utilidade Pública / Energia Elétrica |
| GGBR4 | GERDAU | PN N1 | Materiais Básicos / Sid Metalurgia |
| CMIG4 | CEMIG | PN N1 | Utilidade Pública / Energia Elétrica |
| BRML3 | BR MALLS PAR | ON NM | Financeiro e Outros |
| BRFS3 | BRF SA | ON NM | Consumo Não Básico / Alimentos Processados |
| TIMP3 | TIM PART S/A | ON NM | Telecomunicação |
| LREN3 | LOJAS RENNER | ON NM | Consumo Cíclico / Comércio |

| Código | Empresa | Tipo | Setor e Segmento |
|---------------|----------------|-------------|---|
| WEGE3 | WEG | ON NM | Bens Industriais / Máquinas e Equipamentos |
| EMBR3 | EMBRAER | ON NM | Bens Industriais / Materiais Transporte |
| GOAU4 | GERDAU MET | PN N1 | Materiais Básicos / Sid Metalurgia |
| BBSE3 | BBSEGURIDADE | ON NM | Financeiro e Outros |
| CSNA3 | SID NACIONAL | ON | Materiais Básicos / Sid Metalurgia |
| KLBN11 | KLABIN S/A | UNT N2 | Materiais Básicos / Madeira e Papel |
| RENT3 | LOCALIZA | ON NM | Diversos |
| USIM5 | USIMINAS | PNA N1 | Materiais Básicos / Sid Metalurgia |
| TOTS3 | TOTVS | ON NM | Tec.Informação / Programas Serviços |
| VIVT4 | TELEF BRASIL | PN | Telecomunicação |
| HYPE3 | HYPERA | ON NM | Saúde/Comércio Distribuição. |
| TRPL4 | TRAN PAULIST | PN N1 | Utilidade Pública / Energia Elétrica |
| MRFG3 | MARFRIG | ON NM | Consumo Não Básico / Alimentos Processados |
| ELET3 | ELETRONBRAS | ON N1 | Utilidade Pública / Energia Elétrica |
| SANB11 | SANTANDER BR | UNT | Financeiro e Outros |
| SBSP3 | SABESP | ON NM | Utilidade Pública / Água Saneamento |
| LCAM3 | LOCAMERICA | ON NM | Diversos |
| LIGT3 | LIGHT S/A | ON NM | Utilidade Pública / Energia Elétrica |
| FLRY3 | FLEURY | ON NM | Saúde/Serviço médico |
| YDUQ3 | YDUQS PART | ON NM | Diversos |
| MRVE3 | MRV | ON NM | Consumo Cíclico/Construção Civil |
| ENBR3 | ENERGIAS BR | ON NM | Utilidade Pública / Energia Elétrica |
| JHSF3 | JHSF PART | ON NM | Consumo Cíclico/Construção Civil |
| QUAL3 | QUALICORP | ON NM | Saúde/Serviço médico |
| CYRE3 | CYRELA REALT | ON NM | Consumo Cíclico/Construção Civil |
| SULA11 | SUL AMERICA | UNT N2 | Financeiro e Outros |
| MULT3 | MULTIPLAN | ON N2 | Financeiro e Outros |
| DTEX3 | DURATEX | ON NM | Materiais Básicos / Madeira e Papel |
| BRKM5 | BRASKEM | PNA N1 | Materiais Básicos / Químicos |
| BEEF3 | MINERVA | ON NM | Consumo N Básico / Alimentos Processados |
| EGIE3 | ENGIE BRASIL | ON NM | Utilidade Pública / Energia Elétrica |
| ELET6 | ELETRONBRAS | PNB N1 | Utilidade Pública / Energia Elétrica |
| BRAP4 | BRADSPAR | PN N1 | Materiais Básicos / Mineração |
| TAEE11 | TAESA | UNT N2 | Utilidade Pública / Energia Elétrica |
| RADL3 | RAIADROGASIL | ON NM | Saúde/Comércio Distribuição. |
| BTOW3 | B2W DIGITAL | ON NM | Consumo Cíclico / Comércio |
| CPFE3 | CPFL ENERGIA | ON NM | Utilidade Pública / Energia Elétrica |
| CESP6 | CESP | PNB N1 | Utilidade Pública / Energia Elétrica |
| RAPT4 | RANDON PART | PN N1 | Bens Industriais / Materiais Transporte |
| ECOR3 | ECORODOVIAS | ON NM | Bens Industriais/Transporte |
| ALPA4 | ALPARGATAS | PN N1 | Consumo Cíclico / Tecido Vestuário Calçados |
| LINX3 | LINX | ON NM | Tec.Informação / Programas Serviços |
| CVCB3 | CVC BRASIL | ON NM | Consumo Cíclico/Viagens e Lazer |

| Código | Empresa | Tipo | Setor e Segmento |
|--------|--------------|--------|--|
| GOLL4 | GOL | PN N2 | Bens Industriais/Transporte |
| PRI03 | PETRORIO | ON NM | Petróleo, Gás e Biocombustíveis |
| CSAN3 | COSAN | ON NM | Petróleo, Gás e Biocombustíveis |
| CPLE6 | COPEL | PNB N1 | Utilidade Pública / Energia Elétrica |
| EZTC3 | EZTEC | ON NM | Consumo Cíclico/Construção Civil |
| PSSA3 | PORTO SEGURO | ON NM | Financeiro e Outros |
| IGTA3 | IGUATEMI | ON NM | Financeiro e Outros |
| MDIA3 | M.DIASBRANCO | ON NM | Consumo Não Básico / Alimentos Processados |
| CSMG3 | COPASA | ON NM | Utilidade Pública / Água Saneamento |

FONTES: O autor (2021)

3.1.2 Índices de mercado

Os índices de ações brasileiras selecionados para a execução do trabalho foram: Ibovespa (IBOV), Índice Brasil 100 (IBrX-100), Índice Small Caps (SMLL11) e Índice MidLarge Caps (MLCX). A Tabela 1 apresenta a representatividade perante o valor de mercado da B3 e seus respectivos valores de mercado.

TABELA 1 – Índices de ações.

| Índice | Valor de mercado (R\$ trilhões) | % do valor de mercado B3 |
|---------------|---------------------------------|--------------------------|
| Ibovespa | 3,33 | 82% |
| IBrX-100 | 3,55 | 88% |
| Small Caps | 0,49 | 12% |
| MidLarge Caps | 3,23 | 80% |

FONTES: B3 (2020)

3.1.3 Palavras relacionadas ao sentimento dos investidores.

Para a elaboração da amostra de palavras associadas as expectativas sobre as condições econômicas do país e o sentimento dos investidores, selecionamos, primeiramente, os 118 termos financeiros em inglês propostos no trabalho de Da et al. (2015) para a construção do índice FEARS (*Financial and Economic Attitudes Revealed by Search*). Por meio da ferramenta de tradução do Google, Google Translate, as palavras foram traduzidas para o português, conforme apresentadas no Apêndice 1.

Parte dos termos utilizados por Da et al. (2015) são relacionados ao comportamento específico do norte-americano, sendo necessária a exclusão desses termos de nossa amostra, por exemplo: “401K” e “Chapter 7”. Excluímos também termos que não tiveram dados de frequência de buscas apresentados pelo Google Trends, em razão do baixo nível de pesquisas ao longo do período analisado. Com o objetivo de complementar a seleção de palavras, optou-se por incluir termos relacionadas ao comportamento econômico do brasileiro, como “Serasa”, “Seguro desemprego”, “Bolsa Família” e entre outros.

Dessa forma, de um total de 118 termos propostos por Da et al. (2015), 45 palavras foram excluídas da seleção e 34 foram adicionadas, totalizando uma listagem de 107 termos, conforme apresentada no Apêndice B.

3.2 MÉTRICAS E COLETA DOS DADOS

Essa seção tem o objetivo de detalhar a coleta dos dados e explicar o procedimento de cálculo das principais métricas utilizadas para o desenvolvimento da pesquisa. O Quadro 2 apresenta o resumo das principais métricas.

QUADRO 2 – Resumo das principais métricas utilizadas

| Variável | Definição | Fórmula |
|--------------|---|---|
| SVI | Evolução do índice de buscas no Google Trends | SVt |
| Δ SVI | Evolução das buscas em relação à semana anterior | $\Delta SVI_{i,t} = \ln(SVI_{i,t}) - \ln(SVI_{i,t-1})$ |
| ASVI | Evolução das buscas em relação a mediana de 4 semanas | $ASVI_{i,t} = \ln(SVI_{i,t}) - \ln med(SVI_{i,t-1}, SVI_{i,t-4})$ |
| Retorno | Retorno semanal | $Retorno_{i,t} = \frac{Preço_{i,t}}{Preço_{i,t-1}}$ |
| Abn_Ret | Retornos anormais | $R_{pt} - R_{ft} = \alpha + \beta_m(R_{mt} - R_{ft}) + \beta_sSMB_t + \beta_hHML_t + \varepsilon_t$ |

| Variável | Definição | Fórmula |
|-----------------|--------------------------------|--|
| Δ Volume | Volume semanal | $\Delta \text{volume}_{i,t} = \frac{\text{Volume}_{i,t}}{\text{Volume}_{i,t-1}}$ |
| Abn_Vol | Volumes anormais de negociação | $\text{Abn_Vol}_{i,t} = \ln(\text{vol}_{i,t}) - \ln \text{med}(\text{vol}_{i,t-1}, \text{vol}_{i,t-4})$ |
| Volatilidade | Volatilidade semanal | $\text{volatilidade}_{i,t} = \sqrt{\frac{(\sum_{d=1}^n r_d^2)}{n-1}} * \sqrt{n}$ |
| FEARS Consol | Índice de sentimento | $\text{FEARS_CONSOL}_t = \sum_{i=1}^{30} R_+^i(\Delta \text{MSVI}_{w,t}) - \sum_{i=1}^{30} R_-^i(\Delta \text{MSVI}_{w,t})$ |
| FEARS P | Índice de sentimento Positivo | $\text{FEARS_P}_t = \sum_{i=1}^{30} R_+^i(\Delta \text{MSVI}_{w,t})$ |
| FEARS N | Índice de sentimento negativo | $\text{FEARS_N}_t = \sum_{i=1}^{30} R_-^i(\Delta \text{MSVI}_{w,t})$ |

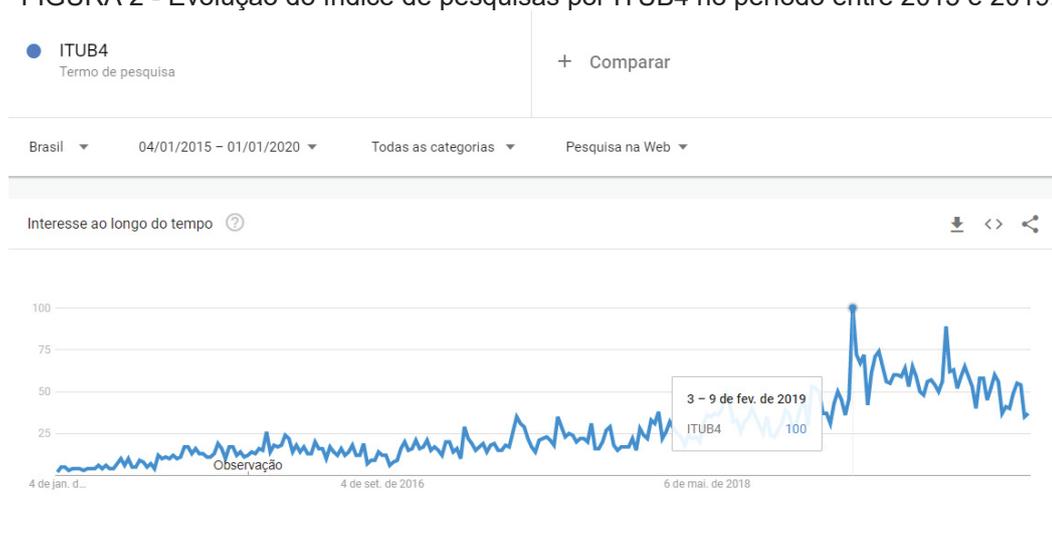
FONTE: O autor (2021)

3.2.1 Frequência de buscas no Google

O Google Trends fornece a evolução dos índices de buscas de uma determinada palavra chave ao longo de um período de tempo, sendo os dados normalizados e apresentados em uma escala 0 – 100. Para calcular esta frequência relativa, cada valor nominal para um intervalo específico é dividido pelo valor máximo no mesmo período. Quando o índice for igual ao valor 100, significa que esse período foi o de maior popularidade dentro do intervalo pesquisado.

Para exemplificar, a Figura 2 apresenta a evolução do índice de pesquisas por “ITUB4” (*ticker* de negociação das ações PN do Itaú) ao longo do período entre 2015 e 2019, sendo a semana a primeira semana de fevereiro de 2019 a de maior interesse dos usuários do Google no Brasil.

FIGURA 2 - Evolução do índice de pesquisas por ITUB4 no período entre 2015 e 2019.



Fonte: Google Trends (2021)

É importante salientar que cada registro gerado pela ferramenta varia de acordo com o tamanho do período selecionado, por exemplo: caso o intervalo selecionado seja de 5 anos, a evolução da frequência de buscas é apresentada em uma periodicidade semanal, enquanto que para intervalos superiores a 5 anos, a periodicidade se torna mensal.

Para medir a atenção e o sentimento dos investidores nas hipóteses verificadas, o presente trabalho utilizou uma periodicidade semanal da evolução da frequência de buscas de cinco principais categorias de palavras-chave, conforme necessidade das hipóteses apresentadas anteriormente, sendo elas: (i) código de negociação das empresas selecionadas; (ii) nome das empresas selecionadas; (iii) nome das empresas selecionadas + palavra “ações” (iv) nome dos índices de mercado e (v) palavras relacionadas ao sentimento dos investidores. A relação das palavras analisadas é apresentada no Quadro 3. Os dados foram coletados via programação em linguagem Python, por meio da inserção automática dos termos de busca na barra de busca do Google Trends e posterior download dos SVIs consultados em formato .CSV.

QUADRO 3 – Termos de busca analisados por categoria.

| Categoria | Termo buscado no Google |
|----------------------------------|---|
| Ticker | ITUB4; ITSA4; ABEV3 ; BBDC4 ; PETR3 ; VALE3 ; B3SA3 ; VVAR3 ; JBSS3; BBAS3 ; BBDC3; CCRO3; CIEL3; UGPA3; EQTL3 ; GGBR4; CMIG4; BRML3; BRFS3; PETR4 ; TIMP3; WEGE3; EMBR3; GOAU4; BBSE3; CSNA3; KLBN11; RENT3; USIM5; TOTS3 ; VIVT4 ; HYPE3 ; TRPL4 ; MRFG3 ; ELET3 ; SANB11 ; LREN3 ; SBSP3 ; LCAM3 ; LIGT3 ; FLRY3 ; YDUQ3 ; MRVE3 ; ENBR3 ; JHSF3 ; QUAL3 ; CYRE3 ; SULA11 ; MULT3 ; DTEX3 ; BRKM5 ; BEEF3 ; EGIE3 ; ENGI11 ; BRAP4 ; TAAE11; RADL3 ; CPFE3 ; CESP6 ; RAPT4 ; ECOR3 ; ALPA4 ; LINX3 ; CVCB3 ; GOLL4 ; PRIO3 ; CSAN3 ; CPLE6; EZTC3 ; PSSA3 ; IGTA3 ; MDIA3 ; CSMG3 ; BTOW3 |
| Empresa Empresa + “ações” | Itaú; Itausa; Ambev; Bradesco; BR Distribuidora; Vale; B3; Via Varejo; JBS; Banco do Brasil; CCR ; Cielo ; Ultrapar ; Equatorial ; Gerdau ; Cemig ; Br Malls; BRF; Petrobras; Tim; Weg; Embraer; Metalurgica Gerdau; BB Seguridade; CSN; Klabin; Localiza; Usiminas; Totvs; Telefonica; Hypera; CTEEP; Marfrig; Eletrobras; Santander ; Lojas Renner ; Sabesp ; Locamerica ; Light S/A ; Fleury ; YDUQS ; MRV ; Energias Br ; JHSF ; Qualicorp ; Cyrela ; Sulamerica ; Multiplan ; Duratex ; Braskem ; Minerva ; Engie Brasil ; Eletrobras ; Bradespar ; Taesa ; Azul ; Drogasil ; CPFL ; Cesp ; Randon ; Ecorodovias ; Alpargatas ; Pão de açúcar ; Linx ; CVC ; Gol linhas aéreas ; Petrorio ; Cosan ; Copel ; Eztec ; Porto Seguro ; Iguatemi ; M Dias Branco ; Copasa ; B2W |
| Índice de mercado | Ibovespa, IBrX, Small Caps e MidLarge Caps |
| Palavras | Abrir Falência ; Banco Comunitário ; Banco Cooperativo ; Benefício ; Benefícios ; Calculadora De Inflação ; Calculadora De Poupança ; Capitalização ; Caridade ; Comprar Ouro ; Conta Poupança ; Contabilidade De Custos ; Crédito Baixo ; Credor ; Crise ; Crise Financeira ; Custo De Vida ; Declarar Falência ; Déficit ; Déficit Brasileiro ; Déficit Comercial ; Déficit Orçamentário ; Demitido ; Depressão ; Desempregado ; Desemprego ; Despesa ; Despesas ; Doação ; Economia ; Empreendedorismo ; Empresarial ; Empresário ; Empréstimo De Capital De Giro ; Falência ; Falido ; Fundo De Patrimônio ; Grande Depressão ; Impostos ; Inflação ; Inflação Brasileira ; Lei De Falências ; Linha Da Pobreza ; Liquidação ; Lucro ; Margem ; Margem De Lucro ; Negócios De Sucesso ; O Déficit ; Ouro ; Parceria Comercial ; Patrimônio ; Patrimônio Líquido ; PIB ; Pobreza ; Pobreza Brasileira ; Pobreza Mundial ; Poupança ; Poupança Brasileira ; Preço Do Ouro ; Private Equity ; Recessão ; Rentável ; Riqueza ; Seguro desemprego ; Seguro Desemprego ; Seguro Saúde ; Tarifa ; Tarifas ; Taxa De Inflação ; Taxa De Pobreza ; Taxas De Inflação ; A Crise ; Salario Mínimo ; Dividas ; Selic ; desigualdade ; hipoteca ; aluguel ; conta atrasada ; IPO ; Banco ; Aposentadoria ; Auxílio ; Bolsa Família ; CDI ; Cesta básica ; Debito ; Divida ; Empréstimo ; Desconto ; Financiamento ; Informalidade ; INSS ; Juros ; Negativo No Banco ; Orçamento Familiar ; Plano De Saúde ; Serasa ; Subsídio ; Taxa de Financiamento ; FGTS ; Taxa de corretagem ; Dólar ; investir em ações ; Renda Fixa ; Tesouro direto |

FONTE: O autor (2021)

A opção por incluir a frequência de busca em códigos de negociação e em nomes da empresa tem como objetivo analisar qual dos dados possui poder preditivo em relação aos dados financeiros do mercado brasileiro. De um lado, Da et al. (2011) e Joseph et al. (2011) abordaram dois principais problemas em utilizar o nome da empresa como indicador de atenção, em primeiro lugar, os investidores podem pesquisar o nome da empresa por razões não relacionadas a decisão de investimento e, em segundo lugar, os investidores podem pesquisar a mesma empresa usando várias variações de seu nome, o que prejudicaria a análise da evolução das pesquisas.

Por outro lado, os trabalhos de Takeda e Wakao (2014), Vlastakis e Markellos (2012) e Latoeiro (2012) defendem que a utilização de *tickers* pode excluir investidores particulares e menos sofisticados, alguns *tickers* possuem nomes genéricos e, geralmente, o nome da empresa é mais buscado que o *ticker*, reduzindo possíveis ruídos.

É importante ressaltar que existe uma limitação sobre entender a real motivação de cada pesquisa realizada no Google. Outro fator limitante é de que as pesquisas não possibilitam entender o perfil dos usuários e investidores, dessa forma, a amostra de dados pode se tornar enviesada para pesquisas realizadas apenas por investidores Pessoa Física.

A partir dos dados coletados via Google Trends, duas variáveis foram parametrizadas para execução dos modelos de regressão, discutidos nos próximos tópicos. Primeiramente, a variável ΔSVI é mensurada com base na evolução da frequência em relação à semana anterior, servindo para identificar se o termo procurado obteve um acréscimo ou decréscimo semanal em sua atenção. O cálculo da variável é representado na Equação 3.1:

$$\Delta SVI_{i,t} = \ln(SVI_{i,t}) - \ln(SVI_{i,t-1}) \quad (3.1)$$

Onde,

$\Delta SVI_{i,t}$: Variação da frequência de buscas do termo i , na semana t em relação a semana $t-1$.

$SVI_{i,t}$: Frequência de buscas do termo i , na semana t

Adicionalmente, seguindo Da et al. (2011), a variável ASVI é calculada pelo log natural do SVI durante a semana menos a mediana do SVI de 4 semanas anteriores. Teoricamente, o SVI da mediana captura níveis normais de atenção ao longo do período, ressaltando saltos\choques recentes ou anormais. A equação 3.2 representa a métrica de cálculo da frequência de busca anormal.

$$ASVI_{i,t} = \ln(SVI_{i,t}) - \ln \text{mediana}(SVI_{i,t-1}, \dots, SVI_{i,t-4}) \quad (3.2)$$

Onde,

$ASVI_{i,t}$: SVI anormal do termo i , na semana t .

$SVI_{i,t}$: Frequência de buscas do termo i , na semana t

O Quadro 4 apresenta os nomes das variáveis ASVI e Δ SVI empregados conforme categoria do termo de busca analisado.

QUADRO 4 – Variável conforme categoria do termo de busca

| Variável | Categoria do termo de busca |
|----------------------|-----------------------------|
| ASVI_ticker | Ticker |
| ASVI_empresa | Nome da Empresa |
| ASVI_ações | Empresa + “ações” |
| ASVI_índice | Índice de mercado |
| Δ SVI_ticker | Ticker |
| Δ SVI_empresa | Nome da Empresa |
| Δ SVI_ações | Empresa + “ações” |
| Δ SVI_índice | Índice de mercado |

FONTE: O autor (2021)

Ressaltando, que todas as medidas de probabilidade de busca foram convertidas em logaritmos naturais, conforme procedimentos adotados por Da et al. (2011), Vozlyublennaia (2014) e Bank et al. (2011).

3.2.2 Indicadores financeiros

Para avaliar os efeitos da atenção e do sentimento dos investidores no mercado acionário brasileiro, selecionamos indicadores de desempenho de mercado, como: retornos anormais, retornos semanais, volumes anormais de negociação, variação do volume negociado semanal e a volatilidade dos ativos. Esses indicadores foram utilizados como variáveis dependentes ou independentes nos modelos econométricos apresentados posteriormente.

Os principais dados foram extraídos do site Yahoo! Finance, em frequência semanal com o fechamento dos dados na sexta-feira de cada semana, considerando o conjunto de ações apresentadas no capítulo 3.1 e para os principais *benchmarks* de mercado.

3.2.2.1 Retornos anormais e semanais

Campbell et al. (1997) sugeriram a criação de uma métrica de retorno anormal dos títulos, definindo-a como a diferença entre o retorno esperado (normal) e o retorno

atual do portfólio, conforme representando na equação 3.3. Sua principal funcionalidade é indicar a de retornos incomuns de títulos, durante determinado intervalo de tempo, e que não podem ser explicadas pela taxa de retorno do mercado.

$$\text{Abn_Ret}_{i,t} = R_{i,t} - E(R_{i,t}) \quad (3.3)$$

Onde,

$\text{Abn_Ret}_{i,t}$: Retorno anormal do ativo i, na semana t

$R_{i,t}$: Retorno semanal do ativo i, na semana t

$E(R_{i,t})$: Retorno esperado do ativo i, para a semana t

Dentre os modelos de precificação do valor esperado, destacamos: Capital Asset Pricing Model (CAPM), Intertemporal Capital Asset Pricing Model (ICAPM), Downside Capital Asset Pricing Model (D-CAPM), o modelo de 3 Fatores de Fama & French e o modelo de 4 fatores proposto por Carhart (1997).

Optamos por utilizar o modelo de precificação de 3 fatores de Fama & French (1993), pela disponibilidade dos dados de cada fator no site do NEFIN (Núcleo de Pesquisa em Economia FEA-USP) e, em razão, da validade e robustez do modelo no mercado acionario brasileiro terem sido comprovadas na literatura nacional, como nos artigos de Júnior (2011) e Malaga e Securato (2003). Fama e French (1993) expandiram o modelo CAPM ao incorporar as métricas de valor de mercado e fator de tamanho da empresa.

Isto posto, no presente trabalho, os retornos anormais são medidos pelo índice alfa de Jensen, estimados pela seguinte regressão 3.4:

$$R_{pt} - R_{ft} = \alpha + \beta_m(R_{mt} - R_{ft}) + \beta_s \text{SMB}_t + \beta_h \text{HML}_t + \varepsilon_t \quad (3.4)$$

Onde,

R_{pt} : retorno do ativo ou portfólio p no período t

R_{ft} : é o retorno do ativo livre de risco no período

$R_{mt} - R_{ft}$: o retorno em excesso em relação ao mercado.

SMB_t : diferença entre os retornos de um portfólio diversificado com ações de baixo valor de Mercado e alto valor de mercado

HML_t : diferença de retorno entre uma carteira com o indicador “Book to Market” alto e baixo.

UMD_t : a diferença de retorno entre uma carteira de ações com alto retorno no ano passado e uma carteira de ações com baixo retorno no ano passado.

ε_t : Erro na estimativa após a regressão.

α : é a taxa de retorno anormal do título (alfa de Jensen).

Ressaltando que utilizamos os retornos semanais ajustados pelos eventos de dividendos e desdobramentos e que os índices foram estimados com base em uma janela móvel com observações dos 12 meses anteriores.

Já o retorno semanal, utilizado como variável dependente em algumas regressões estimadas no trabalho, é medido pela razão entre o valor de fechamento do título na semana t e o valor de fechamento na semana t-1, conforme apresentado na equação 3.5:

$$\text{Retorno}_{i,t} = \frac{\text{Preço}_{i,t}}{\text{Preço}_{i,t-1}} \quad (3.5)$$

Onde,

$\text{Retorno}_{i,t}$: Retorno do ativo i na semana t

$\text{Preço}_{i,t}$:Cotação do ativo i na semana t

3.2.2.2 Volumes anormais e semanais

Outro parâmetro utilizado para análise da relação entre Google Trends e o mercado brasileiro de ações está relacionado aos volumes anormais de negociação de cada título. Ele é medido como a diferença entre o volume semanal de negociações do ativo e a mediana do volume de negociação de uma janela temporal específica. O volume anormal é apresentado na equação 3.6 abaixo:

$$\text{Abn_Vol}_{i,t} = \ln(\text{volume}_{i,t}) - \ln \text{mediana}(\text{volume}_{i,t-1}, \dots, \text{volume}_{i,t-4}) \quad (3.6)$$

Onde,

Abn_Voli, t : Volume anormal do ativo i, na semana t

$volume_{i,t}$: Volume de negociação do ativo i, na semana t

Já a variação do volume semanal, utilizado como variável dependente em algumas regressões estimadas no trabalho, é medida pela razão entre o volume negociado na semana t e o valor negociado na semana t -1, conforme apresentado na equação 3.7. Essa métrica, por exemplo, foi utilizada nos trabalhos de Joseph et al. (2011) e Latoeiro (2012).

$$\Delta volume_{i,t} = \frac{Volume_{i,t}}{Volume_{i,t-1}} \quad (3.7)$$

Onde,

$\Delta volume_{i,t}$ = variação do volume semanal

$volume_{i,t}$: Volume de negociação do ativo i, na semana t.

3.2.2.3 Volatilidade semanal dos ativos

Volatilidade é uma medida popular para avaliar como o retorno das ações varia conforme o tempo. A intensidade de variação do preço semanal dos ativos pode ser calculada com base desvio padrão dos retornos diários dos ativos e ajustados para uma base semanal , conforme equação abaixo:

$$volatilidade_{i,t} = \sqrt{\frac{\sum_{d=1}^n r_i^2}{n-1}} * \sqrt{n} \quad (3.8)$$

Onde,

$volatilidade_{i,t}$: volatilidade do ativo i, na semana t

r_i^2 : retorno diário do ativo i no dia d.

n : número de observações realizadas na semana t.

3.2.3 Construção do índice de sentimento “FEARS”

Para a construção do índice de sentimento FEARS para o mercado brasileiro, adaptamos a metodologia aplicada por Da et al. (2015). Primeiramente, com base nos dados coletados para a amostra de 107 palavras apresentadas na seção anterior, o índice de frequência de buscas ($\Delta SVI_{i,t}$) é calculado para cada um dos termos semanalmente. Da et al. (2015) observaram que a variação na frequência de busca palavras-chave mostraram padrões de heterocedasticidade e valores extremos. Para mitigar qualquer preocupação com os dados, seguindo os autores, optou-se por fazer um teste de robustez (também conhecido como *winsorizing*), com nível de confiança de 5% e normalizar a série de dados.

O índice é consolidado após o agrupamento dinâmico das frequências de buscas das 30 palavras que, após regressões retroativas *backward rolling*, obtiveram os maiores e menores *t-statistic* na relação entre a evolução semanal da frequência de buscas no Google e o retorno semanal de mercado. É importante ressaltar que Da et al. (2015) não encontraram termos com *t-statistic* significativamente positivos, concentrando-se apenas nos termos com *t-statistic* negativo para a construção de seus índice FEARS. Por outro lado, nossa amostra de dados encontrou relações positivas nessas regressões e, assim como nos trabalhos de Streicher (2019) e Gao et al. (2020), optamos por incluir os maiores *t-statistic* na formação do índice, ao calcular a diferença entre o agrupamento positivo e o agrupamento negativo. Dessa forma, a estrutura do índice pode ser representada conforme equação 3.8:

$$FEARS_CONSOL_t = \sum_{i=1}^{30} R_+^i(\Delta MSVI_{w,t}) - \sum_{i=1}^{30} R_-^i(\Delta MSVI_{w,t}) \quad (3.8)$$

Onde,

$FEARS_CONSOL_t$: Valor do índice na semana t.

$R_+^i(\Delta MSVI_t)$ = variação ajustada no índice semanal de buscas da palavra-chave que teve um ranqueamento da classificação t-estatística com os maiores *t-statistic*.

$R_-^i(\Delta MSVI_t)$ = variação ajustada no índice semanal de buscas da palavra-chave que teve um ranqueamento da classificação t-estatística com os menores *t-statistic*.

Adicionalmente, parametrizamos também o cálculo FEARS_P (Positivo) e o FEARS_N (Negativo), desenvolvidos conforme as palavras-chaves na classificação do maior *t-statistic* e de menor *t-statistic*, respectivamente, sendo calculados conforme equações 3.9 e 3.10. Em outras palavras, o índice negativo é formado pelo conjunto de palavras com a maior correlação negativa com o retorno do Ibovespa, enquanto o índice positivo é associado com as palavras com maior correlação positiva.

$$FEARS_P_t = \sum_{i=1}^{30} R_+^i(\Delta MSVI_{w,t}) \quad (3.9)$$

$$FEARS_N_t = \sum_{i=1}^{30} R_-^i(\Delta MSVI_{w,t}) \quad (3.10)$$

A atualização da composição do agrupamento das 30 palavras do índice de sentimento foi realizada a cada 6 meses, conforme períodos apresentados no Quadro 5 abaixo. Como este procedimento requer 180 dias de dados iniciais para calcular o índice FEARS, nosso período de análise começa em julho de 2015.

QUADRO 5 – Datas de atualização do agrupamento de palavras.

| # | Data de vigência da seleção de palavras |
|---|---|
| 1 | Julho de 2015 a dezembro de 2015 |
| 2 | Janeiro de 2016 a junho de 2016 |
| 3 | Julho de 2016 a dezembro de 2016 |
| 4 | Janeiro de 2017 a junho de 2017 |
| 5 | Julho de 2017 a dezembro de 2017 |
| 6 | Janeiro de 2018 a junho de 2018 |
| 7 | Julho de 2018 a dezembro de 2018 |
| 8 | Janeiro de 2019 a junho de 2019 |
| 9 | Julho de 2019 a dezembro de 2019 |

FONTE: O autor (2021)

3.3 MÉTODOS E PROCEDIMENTOS ECONOMETRICOS

Na investigação das possíveis relações preditivas entre a atenção e sentimento dos investidores e o desempenho do mercado acionário brasileiro, empregamos os modelos de autoregressão Vetorial de séries temporais (VAR) e dados em painéis, conforme a natureza dos dados analisados. Enquanto que os

dados das séries temporais são compostos por observações de um elemento em intervalos de tempo múltiplos, os dados de painel são constituídos por observações de vários elementos obtidos ao longo de intervalos múltiplos de tempo. Logo, no presente trabalho, series temporais foram utilizadas em razão das análises individuais dos índices de mercado (Ibovespa, IBRX-100, Small Caps e Middle Large Caps) e os dados em painel, em razão do agrupamento das informações dos títulos acionários.

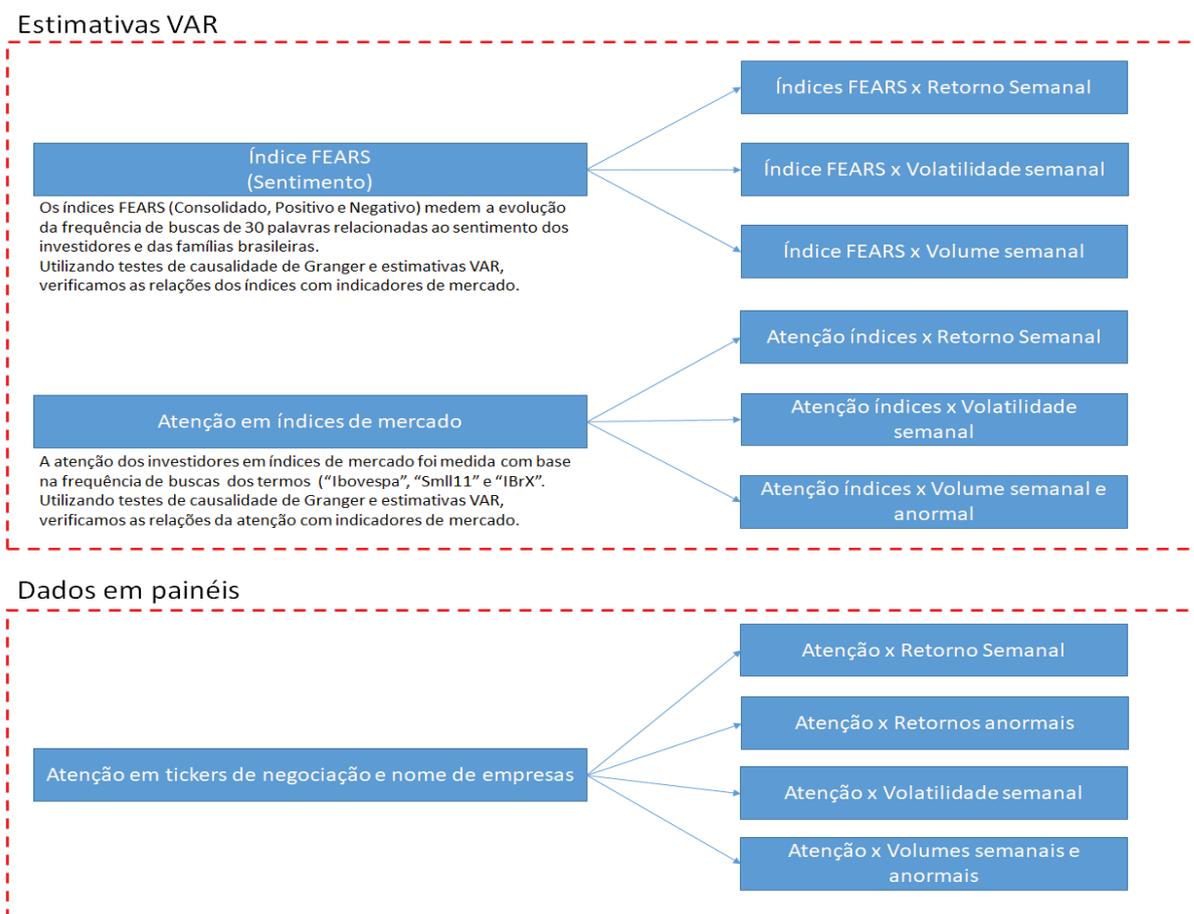
Para facilitar o entendimento das análises e a compreensão dos resultados, classificamos os métodos de regressão utilizados de acordo com as variáveis dependentes e hipóteses associadas. O Quadro 6 e a Figura 3 apresentam essa relação.

QUADRO 6 – Métodos utilizados para testar cada uma das hipóteses.

| Indicador de análise (variável dependente) | Hipótese associada | Método utilizado |
|---|--------------------------------|------------------------------|
| Retornos semanais | H1a - Sentimento FEARS | Autoregressão Vetorial (VAR) |
| Volumes semanais | H1c - Sentimento FEARS | Autoregressão Vetorial (VAR) |
| Volatilidade | H1b - Sentimento FEARS | Autoregressão Vetorial (VAR) |
| Retornos semanais | H2a - Índices de ações | Autoregressão Vetorial (VAR) |
| Volumes semanais | H2b - Índices de ações | Autoregressão Vetorial (VAR) |
| Volumes anormais | H2b - Índices de ações | Autoregressão Vetorial (VAR) |
| Volatilidade | H2c - Índices de ações | Autoregressão Vetorial (VAR) |
| Retornos semanais e anormais | H3a - Ticker ações de empresas | Regressão dados em painéis |
| Retornos semanais e anormais | H4a - Nome de empresas | Regressão dados em painéis |
| Volumes semanais e anormais | H3b - Ticker ações de empresas | Regressão dados em painéis |
| Volumes semanais e anormais | H4b- Nome de empresas | Regressão dados em painéis |

FONTE: O autor (2021)

FIGURA 3 - Métodos utilizados para testar cada uma das hipóteses.



FONTE: O autor (2021)

Por fim, simulamos um modelo prático da utilização das informações extraídas do Google Trends em uma estratégia de investimentos, construída com base em redes neurais artificiais, conforme definição arbitrária do autor. A metodologia de construção dessa estratégia de alocação de capitais e os resultados obtidos comparados ao índice Ibovespa são apresentados no capítulo 7.

4 ÍNDICES FEARS E OS INDICADORES DE MERCADO

Para testar as hipóteses relacionadas aos índices de sentimento FEARS (consolidado, positivo e negativo) e os índices de ações, optou-se por utilizar o modelo de Autoregressão Vetorial (VAR), desenvolvido por Sims (1980). Nesse modelo cada modelo econométrico é estimado por uma regressão MQO de determinada variável dependente em variáveis defasadas de si própria e de outras variáveis, além de considerar todas as variáveis endógenas. Essa metodologia foi utilizada, por exemplo, nos trabalhos de Da et al. (2011), Vozlyublennaia (2014) e Perlin et al. (2016). Adicionalmente, as estimativas VAR foram utilizadas para investigar a causalidade de Granger, verificando se uma série temporal (X) ajuda a prever a outra série (Y), e vice-versa.

Guajurati e Porter (2011) apontam que as principais vantagens do VAR são: (i) é um método simples, pois não é necessário determinar quais variáveis são endógenas e quais são exógenas, uma vez que todas as variáveis em VAR são endógenas; (ii) a estimação das variáveis é simples, visto que o método dos MQO pode ser aplicado em cada modelo separadamente; (iii) as previsões obtidas são, em muitos casos, melhores do que as obtidas com base em modelos mais complexos de equações simultâneas.

A Tabela 2 apresenta as principais estatísticas descritivas para as variáveis das regressões estimadas nessa seção. Destacando que o retorno semanal médio das ações do Ibovespa, e similar aos demais índices acionários, foi em torno de 0,4%, evidenciando que a seleção de dados apresentou viés de altas semanais no período analisado. A média da variação semanal do volume transacionado das ações que compunham o Ibovespa foi positiva em 5,0%, indicando um aumento no fluxo das transações, alinhado com o crescimento de investidores pessoa física na B3. Por fim, os índices de sentimento Positivo e Negativo se mostraram opostos no que tange a média da variação semanal, comportamento em linha com a metodologia aplicada, visto que o índice FEARS positivo é formado pelas palavras com maior correlação histórica com o retorno semanal do Ibovespa e o índice FEARS negativo com as palavras de menor correlação histórica.

TABELA 2 – Estatísticas descritivas – sentimento FEARS e os indicadores de mercado

| | Retorno_Ibov | Retorno_IBrX | Retorno_MLCX | Retorno_SML11 | FEARS_CONSOL |
|-----------|--------------|--------------|--------------|---------------|--------------|
| Média | 0,4% | 0,4% | 0,4% | 0,4% | -1,320 |
| Mediana | 0,5% | 0,5% | 0,5% | 0,7% | -0,759 |
| Máximo | 18,0% | 16,5% | 16,9% | 11,0% | 12,003 |
| Mínimo | -8,2% | -8,1% | -8,2% | -6,9% | -18,879 |
| Desvio P. | 3,0% | 2,9% | 2,9% | 2,8% | 6,174 |

| | FEARS_NEGATIVO | FEARS_POSITIVO | ABN_VOLUME_IBOV | ΔVolume_IBOV | VOLATILIDADE_IBOV |
|-----------|----------------|----------------|-----------------|--------------|-------------------|
| Média | -0,638 | 0,682 | -0,9% | 5,0% | 2,9% |
| Mediana | -0,320 | 0,483 | 0,0% | -0,7% | 2,6% |
| Máximo | 11,098 | 21,524 | 29,1% | 172,0% | 9,4% |
| Mínimo | -14,841 | -15,368 | -45,1% | -58,0% | 0,3% |
| Desvio P. | 5,056 | 5,821 | 11,4% | 33,9% | 1,5% |

Fonte: Dados de pesquisa (2021).

A matriz de correlações das variáveis, apresentada na Tabela 3, indica que as correlações contemporâneas entre o retorno dos índices são superiores a 0,89, ou seja, apresentam uma forte correlação. É possível identificar que apesar de alta, as correlações do retorno do índice SML11 em relação aos outros índices de ações são de menor intensidade, possivelmente pelo SML11 ser formado por empresas com valores de mercado inferiores e perfis de crescimento distintos aos das empresas que compõem os outros índices, além de representarem empresas com maior grau de risco associado.

As correlações entre os índices de sentimento positivo e negativo tiveram baixo nível de intensidade, também apresentando esse comportamento se correlacionadas aos indicadores de mercado. Já o FEARS Consolidado, apresentou correlação positiva de média intensidade com o FEARS Negativo e correlação negativa de média intensidade com o FEARS Positivo, isso pode ser explicado, pois o índice consolidado é uma resultante da diferença entre o índice positivo e do negativo.

TABELA 3 – Matriz de correlações – sentimento FEARS e os indicadores de mercado

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|------------------------|-------|-------|-------|-------|--------|--------|--------|--------|--------|--------|
| (1) Retorno_Ibov | 1,000 | 0,999 | 0,999 | 0,890 | -0,078 | -0,029 | 0,058 | -0,028 | 0,072 | -0,053 |
| (2) Retorno_IBrX | | 1,000 | 0,999 | 0,900 | -0,080 | -0,035 | 0,055 | -0,037 | 0,064 | -0,057 |
| (3) Retorno_MLCX | | | 1,000 | 0,882 | -0,078 | -0,031 | 0,056 | -0,033 | 0,067 | -0,055 |
| (4) Retorno_SML11 | | | | 1,000 | -0,099 | -0,068 | 0,047 | -0,074 | 0,011 | -0,085 |
| (5) FEARS_CONSOL | | | | | 1,000 | 0,477 | -0,646 | 0,044 | -0,188 | -0,002 |
| (6) FEARS_NEGATIVO | | | | | | 1,000 | 0,362 | 0,245 | 0,349 | -0,056 |
| (7) FEARS_POSITIVO | | | | | | | 1,000 | 0,166 | 0,503 | -0,046 |
| (8) ABN_VOLUME_IBOV | | | | | | | | 1,000 | 0,532 | 0,321 |
| (9) ΔVolume_IBOV | | | | | | | | | 1,000 | 0,116 |
| (10) VOLATILIDADE_IBOV | | | | | | | | | | 1,000 |

Fonte: Dados de pesquisa (2021).

É importante ressaltar que Guajurati e Porter (2011) argumentam que todas as variáveis para o modelo de séries temporais VAR devem ser estacionárias, pois caso não sejam, não seria possível generalizar os resultados para outros períodos analisados. Isto posto, para verificar a estacionaridade das séries, foram realizados testes de raiz unitária Dickey-Fuller aumentado (ADF). A hipótese nula afirma que a série segue um processo de raiz unitária, ou seja, é não-estacionária, contra a alternativa de nenhuma raiz unitária, ou seja, série temporal estacionária. Conforme apresentado na Tabela 4, os resultados confirmam que cada série temporal é estacionária.

TABELA 4 – Teste da raiz unitária ADF – Regressão sentimento FEARS e indicadores de mercado.

| Variáveis | Em nível Estatística t | Em nível p-valor | Significância de 1%: |
|----------------------|---------------------------|---------------------|-------------------------|
| Retorno_Ibov | -14,939 | 0.0000 | -3,4619 |
| Retorno_IBrX | -14,904 | 0.0000 | -3,4619 |
| Retorno_MLCX | -14,942 | 0.0000 | -3,4619 |
| Retorno_SML11 | -13,972 | 0.0000 | -3,4619 |
| FEARS_CONSOL | -16,024 | 0.0000 | -3,4619 |
| FEARS_NEGATIVO | -15,384 | 0.0000 | -3,4619 |
| FEARS_POSITIVO | -13,392 | 0.0000 | -3,4619 |
| ABN_VOLUME_IBOV | -10,752 | 0.0000 | -3,4620 |
| Δ Volume_IBOV | -14,539 | 0.0000 | -3,4619 |
| VOLATILIDADE_IBOV | -12,246 | 0.0000 | -3,4620 |

Fonte: Dados de pesquisa (2021).

Complementando a análise de estacionaridade, a Figura 4 apresenta os gráficos de cada série temporal das variáveis utilizadas nessa sessão, totalizando 10 gráficos.

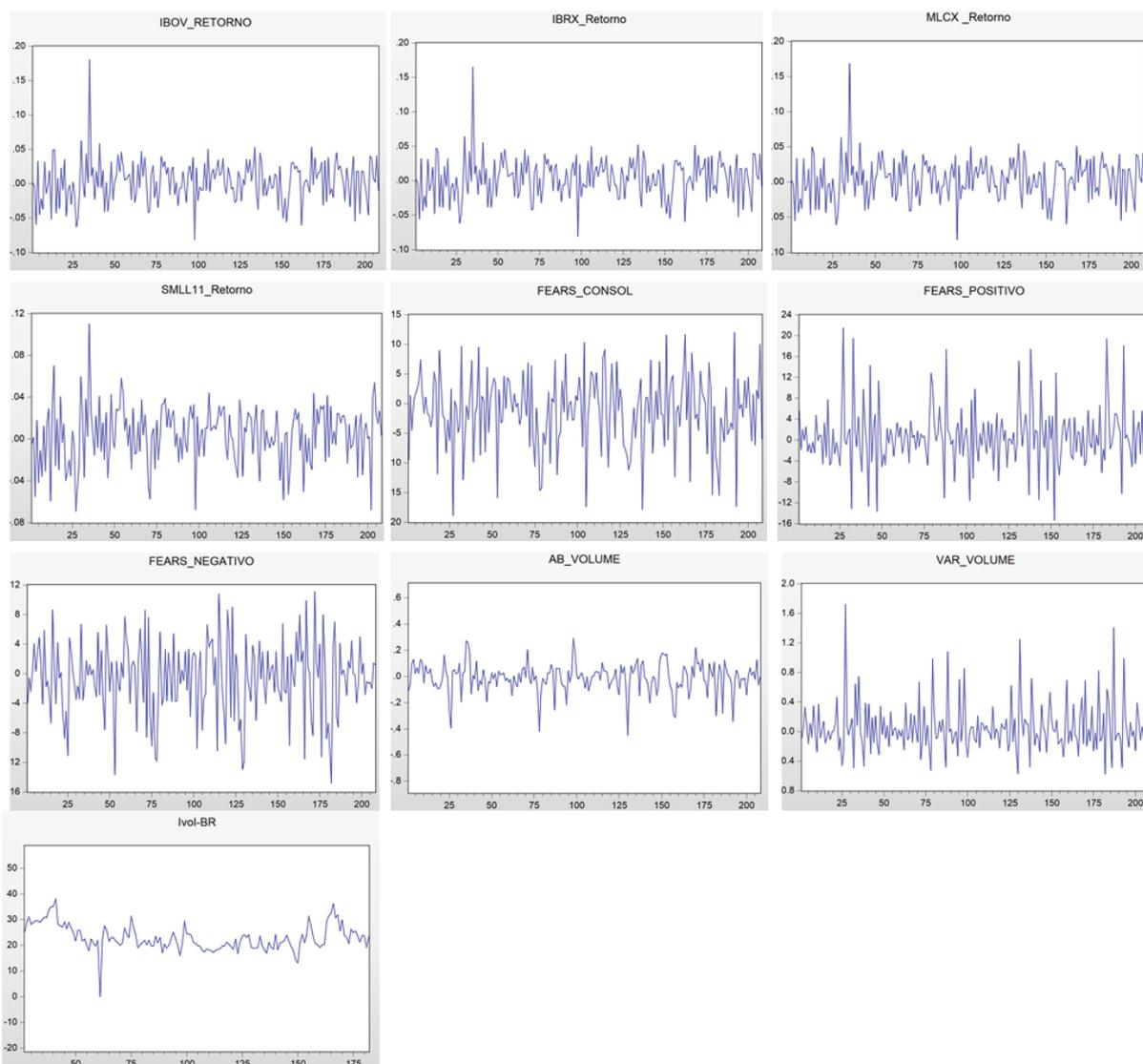


FIGURA 4 – Gráficos das séries temporais das variáveis analisadas nessa seção.
 FONTE: Autor (2021)

Antes de executarmos as causalidades da Granger, exploramos os testes de defasagens ótimas AIC (Akaike Information Criterion), SC (Schwarz) e HQ (Hannan Quinn), em cada uma das relações entre as variáveis analisadas, para uma janela de 12 semanas. Conforme valores sublinhados na Tabela 5, é possível identificar que as seleções se encontram, principalmente, dentro da especificação de até quatro períodos de defasagens. Por esse motivo e seguindo a metodologia de Streicher (2019), optamos por analisar a causalidade de Granger para todos os períodos entre 1 e 4 defasagens.

TABELA 5 – Testes de defasagens – AIC, SC e HQ.

| FEARS_CONSOLIDADO | | | | | | | | | | | | |
|-------------------|-------------------|--------|--------|------------------|--------|--------|----------------|--------|--------|--------------|--------|--------|
| LAGS | Retorno_Ibovespa | | | Retorno_IBRX-100 | | | Retorno_SML11 | | | Retorno_MLCX | | |
| | AIC | SC | HQ | AIC | SC | HQ | AIC | SC | HQ | AIC | SC | HQ |
| 0 | 2,327 | 2,360* | 2,340* | 2,254 | 2,287* | 2,267* | 2,140 | 2,166* | 2,146* | 2,275 | 2,308* | 2,288* |
| 1 | 2,335 | 2,435 | 2,375 | 2,262 | 2,362 | 2,303 | 2,151 | 2,251 | 2,191 | 2,281 | 2,381 | 2,322 |
| 2 | 2,317* | 2,485 | 2,386 | 2,246* | 2,413 | 2,314 | 2,136* | 2,304 | 2,204 | 2,266* | 2,434 | 2,334 |
| 3 | 2,336 | 2,570 | 2,430 | 2,265 | 2,500 | 2,360 | 2,166 | 2,400 | 2,261 | 2,285 | 2,519 | 2,380 |
| 4 | 2,350 | 2,651 | 2,471 | 2,279 | 2,581 | 2,401 | 2,170 | 2,471 | 2,291 | 2,300 | 2,601 | 2,421 |
| 5 | 2,382 | 2,750 | 2,531 | 2,311 | 2,679 | 2,460 | 2,194 | 2,562 | 2,343 | 2,332 | 2,699 | 2,480 |
| 6 | 2,417 | 2,852 | 2,593 | 2,346 | 2,781 | 2,522 | 2,226 | 2,661 | 2,402 | 2,367 | 2,802 | 2,543 |
| 7 | 2,421 | 2,923 | 2,625 | 2,349 | 2,851 | 2,552 | 2,216 | 2,718 | 2,420 | 2,372 | 2,874 | 2,575 |
| 8 | 2,365 | 2,934 | 2,595 | 2,293 | 2,862 | 2,524 | 2,166 | 2,735 | 2,396 | 2,319 | 2,887 | 2,549 |
| 9 | 2,363 | 2,999 | 2,620 | 2,290 | 2,925 | 2,547 | 2,152 | 2,787 | 2,409 | 2,317 | 2,952 | 2,574 |
| 10 | 2,384 | 3,087 | 2,669 | 2,310 | 3,012 | 2,594 | 2,181 | 2,883 | 2,465 | 2,337 | 3,039 | 2,621 |
| 11 | 2,360 | 3,130 | 2,672 | 2,289 | 3,058 | 2,600 | 2,169 | 2,939 | 2,481 | 2,316 | 3,086 | 2,628 |
| 12 | 2,391 | 3,227 | 2,729 | 2,319 | 3,155 | 2,657 | 2,179 | 3,015 | 2,518 | 2,347 | 3,183 | 2,685 |
| FEARS_Positivo | | | | | | | | | | | | |
| LAGS | Retorno_Ibovespa | | | Retorno_IBRX-100 | | | Retorno_SML11 | | | Retorno_MLCX | | |
| | AIC | SC | HQ | AIC | SC | HQ | AIC | SC | HQ | AIC | SC | HQ |
| 0 | 2,230 | 2,263 | 2,243 | 2,158 | 2,191 | 2,171 | 2,040 | 2,073 | 2,053 | 2,178 | 2,211 | 2,191 |
| 1 | 2,139 | 2,239* | 2,179* | 2,065 | 2,165* | 2,105* | 1,968 | 2,068* | 2,008* | 2,083 | 2,183* | 2,124* |
| 2 | 2,115* | 2,283 | 2,183 | 2,042* | 2,209 | 2,110 | 1,949* | 2,117 | 2,017 | 2,0612* | 2,229 | 2,129 |
| 3 | 2,140 | 2,374 | 2,235 | 2,068 | 2,302 | 2,162 | 1,982 | 2,216 | 2,077 | 2,086 | 2,320 | 2,181 |
| 4 | 2,135 | 2,436 | 2,257 | 2,063 | 2,364 | 2,185 | 1,990 | 2,291 | 2,112 | 2,080 | 2,381 | 2,202 |
| 5 | 2,166 | 2,534 | 2,315 | 2,093 | 2,461 | 2,242 | 2,019 | 2,387 | 2,168 | 2,110 | 2,478 | 2,259 |
| 6 | 2,181 | 2,616 | 2,357 | 2,108 | 2,543 | 2,284 | 2,032 | 2,467 | 2,208 | 2,126 | 2,561 | 2,302 |
| 7 | 2,198 | 2,700 | 2,401 | 2,125 | 2,627 | 2,328 | 2,030 | 2,532 | 2,233 | 2,145 | 2,647 | 2,348 |
| 8 | 2,170 | 2,739 | 2,400 | 2,100 | 2,669 | 2,330 | 2,019 | 2,587 | 2,249 | 2,121 | 2,689 | 2,351 |
| 9 | 2,203 | 2,839 | 2,460 | 2,133 | 2,768 | 2,390 | 2,036 | 2,671 | 2,293 | 2,154 | 2,790 | 2,411 |
| 10 | 2,229 | 2,931 | 2,513 | 2,157 | 2,859 | 2,441 | 2,061 | 2,764 | 2,345 | 2,178 | 2,881 | 2,463 |
| 11 | 2,258 | 3,028 | 2,570 | 2,189 | 2,958 | 2,500 | 2,095 | 2,864 | 2,407 | 2,210 | 2,979 | 2,521 |
| 12 | 2,262 | 3,098 | 2,600 | 2,193 | 3,029 | 2,531 | 2,096 | 2,932 | 2,435 | 2,214 | 3,051 | 2,553 |
| FEARS_Negativo | | | | | | | | | | | | |
| LAGS | Retorno_Ibovespa | | | Retorno_IBRX-100 | | | Retorno_SML11 | | | Retorno_MLCX | | |
| | AIC | SC | HQ | AIC | SC | HQ | AIC | SC | HQ | AIC | SC | HQ |
| 0 | 1,945 | 1,969* | 1,949* | 1,862 | 1,896* | 1,876* | 1,741 | 1,774* | 1,754* | 1,883 | 1,916* | 1,896* |
| 1 | 1,944* | 2,045 | 1,985 | 1,860* | 1,970 | 1,910 | 1,733* | 1,834 | 1,774 | 1,880* | 1,992 | 1,932 |
| 2 | 1,965 | 2,132 | 2,033 | 1,890 | 2,058 | 1,958 | 1,751 | 1,918 | 1,819 | 1,913 | 2,080 | 1,981 |
| 3 | 1,996 | 2,230 | 2,091 | 1,921 | 2,155 | 2,016 | 1,780 | 2,014 | 1,875 | 1,944 | 2,179 | 2,039 |
| 4 | 2,004 | 2,305 | 2,126 | 1,928 | 2,229 | 2,050 | 1,791 | 2,092 | 1,913 | 1,950 | 2,251 | 2,072 |
| 5 | 2,017 | 2,385 | 2,166 | 1,940 | 2,308 | 2,089 | 1,795 | 2,163 | 1,944 | 1,962 | 2,330 | 2,111 |
| 6 | 2,051 | 2,486 | 2,227 | 1,975 | 2,409 | 2,151 | 1,832 | 2,267 | 2,008 | 1,997 | 2,432 | 2,173 |
| 7 | 2,041 | 2,543 | 2,244 | 1,963 | 2,464 | 2,166 | 1,812 | 2,314 | 2,015 | 1,987 | 2,488 | 2,190 |
| 8 | 2,016 | 2,585 | 2,246 | 1,939 | 2,508 | 2,170 | 1,793 | 2,362 | 2,024 | 1,964 | 2,533 | 2,194 |
| 9 | 2,034 | 2,669 | 2,291 | 1,957 | 2,593 | 2,214 | 1,808 | 2,443 | 2,065 | 1,982 | 2,618 | 2,239 |
| 10 | 2,046 | 2,749 | 2,331 | 1,968 | 2,671 | 2,253 | 1,828 | 2,530 | 2,112 | 1,993 | 2,696 | 2,278 |
| 11 | 2,068 | 2,837 | 2,379 | 1,992 | 2,761 | 2,304 | 1,852 | 2,622 | 2,164 | 2,017 | 2,786 | 2,328 |
| 12 | 2,067 | 2,903 | 2,406 | 1,993 | 2,829 | 2,331 | 1,853 | 2,689 | 2,192 | 2,016 | 2,853 | 2,355 |
| Δvolume_Ibovespa | | | | | | | | | | | | |
| LAGS | FEARS_CONSOLIDADO | | | FEARS_POSITIVO | | | FEARS_NEGATIVO | | | | | |
| | AIC | SC | HQ | AIC | SC | HQ | AIC | SC | HQ | | | |
| 0 | 7,192 | 7,226 | 7,206 | 6,832 | 6,866 | 6,846 | 6,707 | 6,739* | 6,720 | | | |
| 1 | 7,125 | 7,220* | 7,165 | 6,654 | 6,755* | 6,695 | 6,651 | 6,751 | 6,691 | | | |
| 2 | 7,081 | 7,248 | 7,149 | 6,594* | 6,761 | 6,661* | 6,610 | 6,777 | 6,677* | | | |
| 3 | 7,075 | 7,309 | 7,170 | 6,608 | 6,842 | 6,703 | 6,612 | 6,846 | 6,707 | | | |
| 4 | 7,020 | 7,321 | 7,141* | 6,606 | 6,907 | 6,728 | 6,579 | 6,880 | 6,701 | | | |
| 5 | 7,006* | 7,375 | 7,156 | 6,624 | 6,992 | 6,773 | 6,545* | 6,913 | 6,694 | | | |
| 6 | 7,037 | 7,472 | 7,213 | 6,635 | 7,070 | 6,811 | 6,567 | 7,002 | 6,743 | | | |
| 7 | 7,061 | 7,563 | 7,264 | 6,648 | 7,149 | 6,851 | 6,591 | 7,092 | 6,794 | | | |
| 8 | 7,091 | 7,660 | 7,321 | 6,687 | 7,255 | 6,917 | 6,597 | 7,165 | 6,827 | | | |
| 9 | 7,109 | 7,744 | 7,366 | 6,715 | 7,351 | 6,973 | 6,611 | 7,247 | 6,869 | | | |
| 10 | 7,129 | 7,831 | 7,413 | 6,727 | 7,429 | 7,011 | 6,642 | 7,344 | 6,926 | | | |
| 11 | 7,151 | 7,920 | 7,462 | 6,759 | 7,529 | 7,071 | 6,675 | 7,444 | 6,986 | | | |
| 12 | 7,149 | 7,985 | 7,488 | 6,758 | 7,595 | 7,097 | 6,653 | 7,490 | 6,992 | | | |

| Volatilidade_Ibovespa | | | | | | | | | |
|-----------------------|-------------------|---------------|---------------|----------------|---------------|---------------|----------------|----------------|------------------|
| LAGS | FEARS_CONSOLIDADO | | | FEARS_POSITIVO | | | FEARS_NEGATIVO | | |
| | AIC | SC | HQ | AIC | SC | HQ | AIC | SC | HQ |
| 0 | 1,2495 | 1,253 | 1,251 | 1,209 | 1,212 | 1,210 | 1,239 | 1,242 | 1,240 |
| 1 | 1,1974 | <u>1,207*</u> | <u>1,201*</u> | 1,157 | <u>1,167*</u> | <u>1,161*</u> | 1,180 | <u>11,897*</u> | 1,184 |
| 2 | 1,1962 | 1,213 | 1,203 | <u>1,154*</u> | 1,171 | 1,161 | <u>1,1766*</u> | 1,193 | <u>11,83417*</u> |
| 3 | <u>1,1950*</u> | 1,218 | 1,204 | 1,155 | 1,178 | 1,164 | 1,178 | 1,202 | 1,188 |
| 4 | 1,1972 | 1,227 | 1,209 | 1,156 | 1,186 | 1,168 | 1,181 | 1,211 | 1,193 |
| 5 | 1,2010 | 1,238 | 1,216 | 1,155 | 1,192 | 1,170 | 1,184 | 1,221 | 1,199 |
| 6 | 1,2042 | 1,248 | 1,222 | 1,158 | 1,202 | 1,176 | 1,185 | 1,228 | 1,202 |
| 7 | 1,2071 | 1,257 | 1,227 | 1,159 | 1,209 | 1,180 | 1,188 | 1,238 | 1,208 |
| 8 | 1,2086 | 1,265 | 1,232 | 1,161 | 1,218 | 1,184 | 1,191 | 1,248 | 1,214 |
| 9 | 1,2106 | 1,274 | 1,236 | 1,164 | 1,228 | 1,190 | 1,194 | 1,258 | 1,220 |
| 10 | 1,2137 | 1,284 | 1,242 | 1,165 | 1,235 | 1,194 | 1,196 | 1,266 | 1,224 |
| 11 | 1,2165 | 1,293 | 1,248 | 1,169 | 1,246 | 1,200 | 1,200 | 1,276 | 1,231 |
| 12 | 1,2174 | 1,301 | 1,251 | 1,165 | 1,249 | 1,199 | 1,203 | 1,287 | 1,237 |

Fonte: Dados de pesquisa (2021).

Complementando, para cada análise de regressão VAR complementar, o número de defasagens ótimas foi determinado com base nos testes acima citados, porém nesses casos apenas os resultados com as defasagens ótimas de cada regressão foram apresentados.

Os resultados dos testes de causalidade de Granger e das regressões VAR são apresentados a seguir, divididos em 3 seções: retornos semanais, volumes semanais e volatilidade semanal.

4.1 ÍNDICES FEARS E O INDICADOR DE RETORNO SEMANAL

Primeiramente, por meio do teste de causalidade de Granger, investigamos se o sentimento do investidor (FEARS Consolidado, Positivo e Negativo) possui influência preditiva no retorno semanal dos principais índices de mercado (Ibovespa, IBRX-100, SMLL11 e MLCX) e se variações nos retornos semanais causam mudanças no índice de sentimento do investidor. Essa análise de causalidade, portanto, analisou a relação em ambos os sentidos.

O método de Granger consiste em testar se os coeficientes α nas equações 4.1 e se os coeficientes β na equação 4.2 são conjuntamente diferentes de zero para cada equação. Utilizando os resultados do p-valor resultante, é possível determinar se a hipótese nula de que FEARS não causa Granger do retorno (e vice-versa) pode ser aceita ou não.

$$retorno_{i,t} = \beta_0 + \sum_{k=1}^{lags} \beta FEARS_{i,t-k} + \sum_{k=1}^{lags} \alpha retorno_{i,t-k} \quad (4.1)$$

$$FEARS_{i,t} = \beta_0 + \sum_{k=1}^{lags} \beta FEARS_{i,t-k} + \sum_{k=1}^{lags} \alpha retorno_{i,t-k} \quad (4.2)$$

Onde, a variável FEARS assume os valores do FEARS Consolidado, FEARS Positivo ou FEARS Negativo e a variável 'retorno' assume os valores da variação semanal do retorno do Ibovespa, do IBrX-100, do SMLL11 ou do MLCX. A variável k corresponde ao número de período defasados em relação a variável t.

A Tabela 6 indica os resultados dos testes do p-valor para cada retorno semanal dos índices de mercado e índices FEARS, conforme defasagens (de 1 a 4) utilizadas e apresentadas por coluna, em ambos os sentidos de causalidade.

TABELA 6 – Teste de causalidade de Granger – FEARS e retorno semanal dos índices de mercado
Esta tabela exibe os valores de p para testes de causalidade bilaterais da Granger sobre os índices FEARS (consolidado, positivo e negativo) e os retornos semanais dos índices de ações (Ibovespa, IBRX, SMLL e MLCX). Os resultados se referem à hipótese nula (H0) de ambas as direções relatadas no painel direito e esquerdo, respectivamente. As colunas incluem os resultados sobre as especificações de defasagem de 1 a 4 semanas. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| | | H0: FEARS não causa granger Retorno semanal | | | | H0: Retorno semanal não causa granger FEARS | | | |
|------------|------|---|-----------|----------|----------|---|----------|----------|-----------|
| Defasagens | | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| FEARS_C | Ibov | 0,8044 | 0,1228 | 0,1340 | 0,1332 | 0,0474** | 0,0644* | 0,0527* | 0,074* |
| FEARS_C | IBRX | 0,7790 | 0,1131 | 0,1262 | 0,1210 | 0,0463** | 0,0686* | 0,0621* | 0,0915* |
| FEARS_C | SMLL | 0,4707 | 0,1140 | 0,1433 | 0,0919* | 0,2297 | 0,2098 | 0,3141 | 0,3516 |
| FEARS_C | MLCX | 0,8178 | 0,1209 | 0,1341 | 0,1329 | 0,0392** | 0,0603* | 0,0514* | 0,0782* |
| FEARS_P | Ibov | 0,074* | 0,0206** | 0,0298** | 0,0227** | 0,0051*** | 0,0382** | 0,048** | 0,0064*** |
| FEARS_P | IBRX | 0,0667* | 0,0178** | 0,0273** | 0,0203** | 0,0048*** | 0,037** | 0,0489** | 0,007*** |
| FEARS_P | SMLL | 0,07* | 0,0253** | 0,0423** | 0,0367** | 0,0545* | 0,2467 | 0,3405 | 0,1370 |
| FEARS_P | MLCX | 0,0689* | 0,0192** | 0,0295** | 0,0221** | 0,0041*** | 0,0319** | 0,0414** | 0,0053*** |
| FEARS_N | Ibov | 0,0184** | 0,0506* | 0,0911* | 0,1709 | 0,5969 | 0,8213 | 0,7713 | 0,4522 |
| FEARS_N | IBRX | 0,0141** | 0,0396** | 0,072* | 0,1421 | 0,5821 | 0,8367 | 0,7782 | 0,4289 |
| FEARS_N | SMLL | 0,0028*** | 0,0088*** | 0,0184** | 0,0367** | 0,5618 | 0,6737 | 0,6299 | 0,6341 |
| FEARS_N | MLCX | 0,0176** | 0,0491** | 0,0879* | 0,1685 | 0,6023 | 0,8540 | 0,7923 | 0,4110 |

Fonte: Dados de pesquisa (2021).

Conforme resultados apresentados, em relação ao painel do lado direito, encontramos evidências de que os retornos semanais dos índices Ibovespa, IBRX-100 e MLCX causam Granger do sentimento dos investidores, medido pelos índices FEARS Consolidado e FEARS Positivo, uma vez que várias defasagens indicaram resultados significativos de causalidade. Por exemplo, uma defasagem do Índice FEARS Consolidado no retorno semanal do índice Ibovespa sugere significância a um nível de 5%. Essa relação significa que variações nos índices de preços refletem uma mudança de comportamento dos índices de sentimento nas semanas posteriores.

Em relação ao painel do lado esquerdo, sugere-se que os índices FEARS Positivo e FEARS Negativo podem auxiliar na previsão de resultados de retornos futuros dos índices de mercado analisados, visto a causalidade significativa entre as variáveis. Exemplificando, a relação entre o índice FEARS negativo e o retorno do índice SMLL11, com uma defasagem semanal na análise de causalidade, resultou em um p-valor de 0,0028 com significância estatística de 1%.

Adicionalmente, para a equação 4.1, objeto principal de análise do trabalho proposto, realizamos estimativas do modelo VAR com defasagens ótimas. Por meio dessa análise foi possível avaliar qual o sentido da relação entre as variáveis e a

influência de cada uma das variáveis defasadas. Os coeficientes de regressão, seguidos pelos erros padrões, são apresentados na Tabela 7.

TABELA 7 – Estimativa VAR – Índices FEARS e retorno semanal dos índices de mercado

Esta tabela fornece resultados das estimativas VAR para os índices de ações sobre valores defasados do índice de sentimento FEARS (Consolidado, Positivo e Negativo) e o próprio retorno dos índices de ações. As duas especificações do VAR incluem defasamentos conforme testes de AIC e LR. Os coeficientes reportados são seguidos por erros padrão entre parênteses. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| FEARS_CONSOLIDADO | | | | | FEARS POSITIVO | | | | |
|-------------------|------------------------|-------------------------|--------------------------|-------------------------|----------------|-----------------------|----------------------|----------------------|----------------------|
| Variáveis | Ibovespa Lag = 2 | IBRX100 lag =2 | Sml11 Lag =4 | MLCX Lag =2 | Variáveis | Ibovespa Lag = 2 | IBRX100 Lag = 2 | Sml11 Lag = 2 | MLCX Lag = 2 |
| FEARS t-1 | -0,000102 (0,00035) | -0,000112 (0,00033) | -0,000305 (0,00032) | -9,45E-06 (0,00034) | FEARS_P t-1 | -0,00058 (0,0004) | -0,00057 (0,0004) | -0,0005 (0,0003) | -0,00057 (0,0003) |
| FEARS t-2 | -0,0007** (0,00035) | -0,00070** (0,00033) | -0,000789** (0,00032) | -0,00069** (0,00034) | FEARS_P t-2 | 0,00069* (0,0003) | 0,0006* (0,0003) | 0,00063* (0,0003) | 0,00068* (0,0004) |
| FEARS t-3 | | | 0,00000573 (0,00032) | | FEARS_P t-3 | | | | |
| FEARS t-4 | | | 0,000465 (0,00032) | | FEARS_P t-4 | | | | |
| Retorno t-1 | -0,041842 (0,06957) | -0,040427 (0,06957) | 0,007568 (0,07137) | -0,04195 (0,06960) | Retorno t-1 | -0,01352 (0,0693) | -0,01091 (0,0693) | 0,04321 (0,0693) | -0,01357 (0,0693) |
| Retorno t-2 | 0,107027 (0,07053) | 0,104583 (0,07053) | 0,095032 (0,07073) | 0,101961 (0,07063) | Retorno t-2 | 0,13290* (0,07039) | 0,1324* (0,0703) | 0,12265* (0,0695) | 0,12905* (0,0705) |
| Retorno t-3 | | | 0,01785 (0,07038) | | Retorno t-3 | | | | |
| Retorno t-4 | | | 0,096671 (0,07040) | | Retorno t-4 | | | | |
| Constante | 0,002337 (0,00221) | 0,002327 (0,00213) | 0,002545 (0,00211) | 0,00234 (0,00215) | Constante | 0,00311 (0,0024) | 0,00309 (0,0021) | 0,00325 (0,0020) | 0,00309 (0,0021) |
| Obs | 206 | 206 | 204 | 206 | Obs | 206 | 206 | 206 | 206 |
| R2 Adj. | 0,017979 | 0,018234 | 0,019958 | 0,016956 | R2 Adj. | 0,035256 | 0,036155 | 0,031086 | 0,034812 |
| FEARS NEGATIVO | | | | | | | | | |
| Variáveis | Ibovespa Lag = 1 | IBRX100 Lag = 1 | Sml11 Lag = 1 | MLCX Lag = 1 | | | | | |
| FEARS_N t-1 | -0,0009** (0,00041) | -0,00098** (0,00040) | -0,00113*** (0,00038) | -0,00095** (0,00040) | | | | | |
| Retorno t-1 | -0,047507 (0,06906) | -0,046287 (0,06900) | 0,010668 (0,06862) | -0,048046 (0,06905) | | | | | |
| Constante | 0,003145 (0,00211) | 0,003124 (0,00203) | 0,003234 (0,00192) | 0,003108 (0,00205) | | | | | |
| Obs | 207 | 207 | 207 | 207 | | | | | |
| R2 Adj. | 0,019219 | 0,02124 | 0,034006 | 0,019605 | | | | | |

Fonte: Dados de pesquisa (2021).

Os resultados das estimativas sugerem que o índice FEARS Consolidado e o índice FEARS Negativo possuem um impacto negativo no retorno semanal dos índices de ação Ibovespa, SMLL11, IBrX-100 e MLCX nas semanas posteriores, com coeficientes estatisticamente significativos ao nível de 5%.

Podemos interpretar, por exemplo, que um maior nível de pesquisas em palavras com maior correlação negativa histórica com o retorno das ações, ao formarem o índice FEARS Negativo, podem refletir uma piora no retorno dos principais índices acionários na semana posterior. Essa relação com uma defasagem está em linha com a identificada por Kostopoulos et al. (2020), para o mercado alemão, e Brochado (2019), para o mercado português. Da et al. (2015) apesar de terem identificado uma relação contemporânea positiva entre o índice negativo FEARS e o retorno do mercado americano, para defasagem de uma semana do índice, os autores verificaram uma relação negativa entre as variáveis, ou seja, resultado oposto ao identificado em nossas análises.

Já o índice FEARS positivo, com defasagem de duas semanas, possui um impacto positivo e estatisticamente significativo no retorno semanal dos índices acionário e, assim como identificado por Brochado (2019), o impacto negativo da variável com uma semana de defasagem não foi significativamente estatístico. No caso do Ibovespa, por exemplo, o coeficiente defasado em duas semanas da variável FEARS Positivo foi de 0,00069, significando que um aumento de 1 p.p no índice de sentimento positivo é seguido de um aumento no retorno semanal de 0,069 p.p.

Dessa forma, a construção dos índices de sentimento, tanto positivo quanto negativo, podem servir como uma *proxy* de medição do sentimento dos investidores, uma vez que, nesse caso, antecipou a rentabilidade do mercado com significância estatística, mesmo que com baixo impacto. Tais resultados sustentam a hipótese de que os mercados são afetados pelo *noise trading*, de De Long et al. (1990), demonstrando que o comportamento das ações podem ser influenciadas pelo sentimento dos investidores, afastando seu valor de negociação do seu valor fundamental. Portanto, a hipótese H1a, definida como “um aumento nos índices de sentimento dos investidores reflete um efeito positivo ou negativo no retorno semanal dos índices de mercado”, é aceita

4.2 ÍNDICES FEARS E O INDICADOR DE VOLATILIDADE SEMANAL

Dedica-se essa subseção ao estudo da relação entre a volatilidade do Ibovespa e os indicadores de sentimento FEARS. Conforme apontado por Latoeiro (2012), a volatilidade funciona como um indicador de risco, podendo ser lida como um índice refletor das incertezas de mercado. A partir do momento que a incerteza pode gerar o sentimento de medo e, por consequência, despertar a atenção dos indivíduos, a procura por informações pode ser uma manifestação dessa atenção e desse medo.

Conforme metodologia aplicada na seção 4.1, para o indicador volatilidade, as regressões 4.3 e 4.4 serviram de base para os testes de causalidade de Granger.

$$volatilidade_{i,t} = \beta_0 + \sum_{k=1}^{lags} \beta FEARS_{i,t-k} + \sum_{k=1}^{lags} \alpha volatilidade_{i,t-k} \quad (4.3)$$

$$FEARS_{i,t} = \beta_0 + \sum_{k=1}^{lags} \beta FEARS_{i,t-k} + \sum_{k=1}^{lags} \alpha volatilidade_{i,t-k} \quad (4.4)$$

Onde, a variável FEARS assume os valores do FEARS Consolidado, FEARS positivo e FEARS negativo e a variável volatilidade assume os valores da volatilidade semanal do Ibovespa. Para os demais índices de mercado, não foi possível analisar a variação na volatilidade, devido a limitação de acesso aos conjuntos de dados históricos.

A Tabela 8 apresenta os resultados dos testes do p-valor para cada volatilidade semanal do Ibovespa e os índices de sentimento FEARS, conforme defasagens (de 1 a 4) utilizadas e apresentadas por coluna, em ambos os sentidos de causalidade.

TABELA 8 – Teste de causalidade de Granger – índices FEARS e a volatilidade semanal do Ibovespa
Esta tabela exhibe os valores de p para testes de causalidade bilaterais da Granger sobre os índices FEARS (consolidado, positivo e negativo) e a volatilidade semanal do índice de ações Ibovespa. Os resultados se referem à hipótese nula (H0) de ambas as direções relatadas no painel direito e esquerdo, respectivamente. As colunas incluem os resultados sobre as especificações de defasagem de 1 a 4 semanas. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| H0: FEARS não causa granger volatilidade | | | | | H0: volatilidade não causa granger FEARS | | | | |
|--|------|--------|--------|----------|--|--------|--------|--------|---------|
| Lags | | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| FEARS_C | Ibov | 0.9485 | 0.2149 | 0.0215** | 0.0184** | 0.3914 | 0.6888 | 0.8968 | 0.8611 |
| FEARS_P | Ibov | 0.4019 | 0.5230 | 0.7145 | 0.2552 | 0.2714 | 0.4883 | 0.6728 | 0.5021 |
| FEARS_N | Ibov | 0.2972 | 0.3061 | 0.0205** | 0.0359** | 0.9800 | 0.8507 | 0.9158 | 0.0766* |

Fonte: Dados de pesquisa (2021).

Os testes de causalidade de Granger indicaram que os índices FEARS Negativo e FEARS Consolidado causam Granger da variável volatilidade do índice Ibovespa, com significância de 5%. Estes resultados sugerem que para esse modelo, em termos de previsibilidade da variável volatilidade, utilizar os índices FEARS Consolidado e Negativo é mais indicado que utilizar o indicador FEARS Positivo, visto que este índice não apresentou nenhuma causalidade significativa. Em relação a causalidade inversa, apresentada no lado direito do painel, a volatilidade causou Granger apenas do indicador FEARS Negativo defasado em 4 períodos com uma significância estatística de 10%, ou seja, a volatilidade do mercado causa variações no sentimento dos investidores após 4 semanas.

Adicionalmente, para a equação 4.3 realizamos estimativas do modelo VAR com defasagens ótimas. Os coeficientes seguidos pelos erros padrões são apresentados na Tabela 9.

TABELA 9 – Estimativa VAR - FEARS e a volatilidade semanal

Esta tabela fornece resultados das estimativas VAR para a volatilidade do índice de ações Ibovespa sobre valores defasados do índice de sentimento FEARS (Consolidado, Positivo e Negativo) e a própria volatilidade do Ibovespa. As duas especificações do VAR incluem defasamentos conforme testes de AIC e LR. Os coeficientes reportados são seguidos por erros padrão entre parênteses. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| FEARS CONSOLIDADO | | FEARS POSITIVO | | FEARS NEGATIVO | |
|--------------------------|---------------------------|-----------------------|--------------------------|-----------------------|---------------------------|
| Volatilidade Ibovespa | | Volatilidade Ibovespa | | Volatilidade Ibovespa | |
| Variáveis | Max Lag = 3 | Variáveis | Max Lag = 1 | Variáveis | Max Lag = 4 |
| FEARS t-1 | 0,000066 (0,000162) | FEARS_P t-1 | 0.000145 (0.000173) | FEARS_N t-1 | 0,000245 (0,000197) |
| FEARS t-2 | 0,000306 (0,000163) | FEARS_P t-2 | | FEARS_N t-2 | 0,000274 (0,000198) |
| FEARS t-3 | 0,000441*** (0,000163) | FEARS_P t-3 | | FEARS_N t-3 | 0,000553*** (0,000198) |
| FEARS t-4 | | FEARS_P t-4 | | FEARS_N t-4 | 0,000143 (0,000202) |
| Volatilidade t-1 | 0,112009 (0,069167) | Volatilidade t-1 | 0.157153** (0.069152) | Volatilidade t-1 | 0,112678 (0,071646) |
| Volatilidade t-2 | 0,038672 (0,069305) | Volatilidade t-2 | | Volatilidade t-2 | 0,02745 (0,07) |
| Volatilidade t-3 | 0,133575* (0,06874) | Volatilidade t-3 | | Volatilidade t-3 | 0,155542** (0,069938) |
| Volatilidade t-4 | | Volatilidade t-4 | | Volatilidade t-4 | -0,008709 (0,070136) |
| Constante | 0,021888 (0,003269) | Constante | 0.024408 (0.002255) | Constante | 0,021563 (0,00358) |
| Obs. | 205 | Obs. | 207 | Obs. | 204 |
| R2 ajust. | 0,063254 | R2 ajust. | 0.027243 | R2 ajust. | 0,057605 |

Fonte: Dados de pesquisa (2021).

Os resultados das regressões indicaram que o FEARS consolidado e o FEARS negativo explicam parte da volatilidade do índice acionário Ibovespa, com nível de significância de 10%, conforme relação positiva entre as variáveis. Ou seja, quanto maior for o nível de buscas das palavras que compõem os índices de sentimento em três defasagens, maior será a volatilidade do Ibovespa. É possível dessa forma, indicar que quanto maior a preocupação dos investidores, maiores serão os efeitos na volatilidade do mercado, tal lógica está em linha com a ideia de que volatilidade está associada ao medo. Já em relação aos coeficientes dos índices de sentimento com uma e duas semanas de defasagem, nenhuma significância foi identificada.

Complementando, apesar de terem identificado uma relação positiva entre o índice FEARS Negativo e a volatilidade do mercado, a significância só foi verificada por Da et al. (2015) na relação contemporânea entre as variáveis, ou seja, sem indicar um poder preditivo.

Já em relação ao índice FEARS Positivo, assim como identificado no teste de causalidade de Granger, nenhuma relação significativa foi verificada entre a variável defasada e a volatilidade do Ibovespa.

Dessa forma, é possível sugerir que os resultados identificados suportam a hipótese H1b, definida como “um aumento nos índices de sentimento dos investidores reflete um efeito positivo ou negativo na volatilidade dos índices de mercado”, uma vez que os índices de sentimento podem ser utilizados como previsores da volatilidade de mercado.

4.3 ÍNDICES FEARS E O INDICADOR DE VOLUME SEMANAL

Nessa subseção, foram analisados os impactos das variações dos índices FEARS na variação semanal do volume transacionado do índice Ibovespa. O volume de negociação demonstra o interesse e o apetite de mercado, além ser um indicador relevante de liquidez dos ativos.

Conforme metodologia aplicada nas seções 4.1 e 4.2, para o indicador volume semanal, as regressões 4.5 e 4.6 serviram de base para os testes de causalidade de Granger.

$$\Delta volume_{i,t} = \beta_0 + \sum_{k=1}^{lags} \beta FEARS_{i,t-k} + \sum_{k=1}^{lags} \alpha \Delta volume_{i,t-k} \quad (4.5)$$

$$FEARS_{i,t} = \beta_0 + \sum_{k=1}^{lags} \beta FEARS_{i,t-k} + \sum_{k=1}^{lags} \alpha \Delta volume_{i,t-k} \quad (4.6)$$

Onde, a variável FEARS assume os valores do FEARS Consolidado, FEARS Positivo e FEARS Negativo e a variável Δ volume assume os valores da variação semanal do volume transacionado do Ibovespa. Para os demais índices de mercado, não foi possível analisar essa relação, devido a limitação de acesso aos conjuntos de dados históricos.

A Tabela 10 indica os resultados dos testes do p-valor para cada variação semanal do volume de negociação do Ibovespa e os índices FEARS, conforme defasagens (de 1 a 4) utilizadas e apresentadas por coluna, em ambos os sentidos de causalidade.

TABELA 10 – Teste de causalidade de Granger – FEARS e o volume semanal de negociações
Esta tabela exhibe os valores de p para testes de causalidade bilaterais da Granger sobre os índices FEARS (consolidado, positivo e negativo) e a variação do volume de negociações do índice de ações Ibovespa. Os resultados se referem à hipótese nula (H0) de ambas as direções relatadas no painel direito e esquerdo, respectivamente. As colunas incluem os resultados sobre as especificações de defasagem de 1 a 4 semanas. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| | | H0: FEARS não causa granger volume | | | | H0: volume não causa granger FEARS | | | |
|---------|------|------------------------------------|----------|----------|-----------|------------------------------------|--------|--------|--------|
| Lags | | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| FEARS | Ibov | 0,6784 | 0,0483** | 0,0464** | 0,0004*** | 0,8126 | 0,6019 | 0,6642 | 0,6806 |
| FEARS_P | Ibov | 0,1896 | 0,2107 | 0,3389 | 0,1553 | 0,9067 | 0,8778 | 0,5627 | 0,2735 |
| FEARS_N | Ibov | 0,3915 | 0,0464** | 0,0473** | 0,0024*** | 0,0600** | 0,1387 | 0,1035 | 0,1053 |

Fonte: Dados de pesquisa (2021).

Os testes de causalidade de Granger indicaram que os índices FEARS Negativo e FEARS Consolidado, assim como a análise com o indicador de volatilidade do Ibovespa, causam Granger do volume semanal do índice Ibovespa, com significância de até 1%. Já o índice FEARS Positivo não apresentou nenhuma causalidade significativamente estatística.

Em relação aos resultados da direção oposta de causalidade, representados pelos valores do painel direito da tabela, identificamos significância apenas na relação de Granger entre o volume semanal e o índice FEARS Negativo com uma semana de defasagem, indicando que variações no volume negociado causam alterações nos índices de sentimentos.

Adicionalmente, para a equação 4.5 realizamos estimativas do modelo VAR com defasagens ótimas. Os coeficientes seguidos pelos erros padrões são apresentados na Tabela 11.

TABELA 11 – Estimativa VAR - FEARS e o volume semanal de negociações

Esta tabela fornece resultados das estimativas VAR para a variação semanal do volume do índice de ações Ibovespa sobre valores defasados do índice de sentimento FEARS (Consolidado, Positivo e Negativo) e a próprio volume semanal do Ibovespa. As duas especificações do VAR incluem desfasamentos conforme testes de AIC e LR. Os coeficientes reportados são seguidos por erros padrão entre parênteses. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| FEARS_CONSOLIDADO | | FEARS POSITIVO | | FEARS NEGATIVO | |
|---------------------|----------------------------|---------------------|----------------------------|---------------------|----------------------------|
| Volume_Ibovespa | | Volume_Ibovespa | | Volume_Ibovespa | |
| Variáveis | Max Lag = 4 | Variáveis | Max Lag = 2 | Variáveis | Max Lag = 4 |
| FEARS t-1 | -0,002089 (0,003522) | FEARS_P t-1 | 0,007970* (0,004677) | FEARS_N t-1 | 0,003118 (0,004487) |
| FEARS t-2 | -0,007777** (0,0035570) | FEARS_P t-2 | 0,004804 (0,004694) | FEARS_N t-2 | -0,010664** (0,004488) |
| FEARS t-3 | 0,002288 (0,003609) | FEARS_P t-3 | | FEARS_N t-3 | 0,004631 (0,004557) |
| FEARS t-4 | -0,013193*** (0,00359) | FEARS_P t-4 | | FEARS_N t-4 | -0,396641*** (0,004565) |
| Δ Volume t-1 | -0,402679*** (0,069815) | Δ Volume t-1 | -0,437843*** (0,080805) | Δ Volume t-1 | -0,29686*** (0,072791) |
| Δ Volume t-2 | -0,388723*** (0,073994) | Δ Volume t-2 | -0,281412*** (0,080395) | Δ Volume t-2 | -0,230533*** (0,077001) |
| Δ Volume t-3 | -0,222068*** (0,074198) | Δ Volume t-3 | | Δ Volume t-3 | -0,061688*** (0,075296) |
| Δ Volume t-4 | -0,190589*** (0,069261) | Δ Volume t-4 | | Δ Volume t-4 | 0,086878 (0,072343) |
| Constante | 0,080835 (0,024206) | Constante | 0,078048 (0,02282) | Constante | 0,086878 (0,024573) |
| Obs. | 204 | Obs. | 206 | Obs. | 204 |
| R2 ajust. | 0,227054 | R2 ajust. | 0,132429 | R2 ajust. | 0,210084 |

Fonte: Dados de pesquisa (2021).

Os resultados identificados para a relação entre o volume do Ibovespa e os índices de Sentimento FEARS Negativo e FEARS Consolidado indicaram significância estatística para defasagens de 2 e 4 semanas da variável FEARS, sendo que ambos os coeficientes são negativos, ou seja, uma diminuição em buscas de palavras com maiores correlações negativas históricas com o retorno do Ibovespa prediz um

aumento no volume transacionado no mercado. Isso pode significar que quanto menor a preocupação das pessoas com aspectos financeiros, maior será a operação do mercado nas semanas subsequentes.

Por fim, em relação ao índice de sentimento FEARS Positivo, notou-se significância estatística na variável FEARS com defasagem de uma semana, sendo a relação do coeficiente positiva com a variação de volume. Em tese, quanto maior a procura por palavras positivas, maior será o volume transacionado na semana seguinte. Assim como nas outras análises, os resultados das regressões suportam a hipótese H1c, definida como “um aumento nos índices de sentimento dos investidores reflete um efeito positivo ou negativo no volume negociado dos índices de mercado”.

5 ATENÇÃO EM ÍNDICES DE AÇÕES E OS INDICADORES DE MERCADO

As variações anormais e semanais da atenção dos investidores em índices de ações foram medidas com base nas evoluções dos índices de buscas dos termos referentes aos índices de ações, como: “Ibovespa”, “SMLL11” e “IBrX”. É importante salientar que o volume de pesquisas da palavra “Ibovespa” no Google é relevantemente superior aos volumes de pesquisas dos outros índices, principalmente pela popularidade associada ao termo e por ser o principal *benchmark* do mercado acionário brasileiro. Selecionado inicialmente na amostra de índices, o *Middle and Large Caps* (MLCX) não foi considerado na análise, visto que o Google Trends não retornou nenhuma evolução do índice de busca relevante ao termo para o período analisado.

Conforme metodologia aplicada na seção anterior, para testar as hipóteses relacionadas a variação semanal/anormal da atenção dos investidores em índices de ações e os indicadores de mercado, optou-se por utilizar o modelo de Autoregressão Vetorial (VAR) e os testes de causalidade de Granger. Os indicadores selecionados para análise dos impactos foram: retorno semanal, volatilidade semanal, volume anormal de negociação e volume semanal.

A Tabela 12 apresenta as principais estatísticas descritivas das variáveis utilizadas nos modelos propostos, onde é possível destacar as seguintes observações: (i) o alto desvio padrão na variação semanal do índice de buscas dos índices IBrX-100 e do SMLL11, incluindo oscilações significantes de máximo e mínimo, fato que não foi verificado com o índice Ibovespa (ii) é possível notar uma tendência de crescimento das buscas no Google das palavras “Ibovespa”, “SMLL11” e “IBrX” ao se verificar a positividade da média do indicador de variação das frequências semanais de busca; (iii) os indicadores de mercado, conforme citados na seção anterior, apresentaram viés de valorização no período analisado.

TABELA 12 – Estatísticas descritivas – Atenção dos investidores e os indicadores de mercado

| | Retorno_Ibov | Retorno_IBrX | Retorno_SML11 | Δ SVI_IBOV | Δ SVI_IBrX | Δ SVI_SMLL |
|-----------|--------------|--------------|---------------|-----------------------|----------------------|-------------------|
| Média | 0,4% | 0,4% | 0,4% | 0,01 | 0,01 | 0,01 |
| Mediana | 0,5% | 0,6% | 0,7% | - | - | - |
| Máximo | 18,0% | 16,5% | 11,0% | 1,30 | 4,53 | 3,26 |
| Mínimo | -8,2% | -8,1% | -6,9% | -0,80 | -4,61 | -3,00 |
| Desvio P. | 2,9% | 2,8% | 2,7% | 0,27 | 2,36 | 1,14 |
| | ASV_Ibov | ASV_IBRX | ASV_SMLL | VOLATILIDADE IBOVESPA | Δ Volume_IBOV | |
| Média | 0,03 | 0,08 | -0,05 | 2,8% | 4,4% | |
| Mediana | 0,02 | - | - | 2,5% | -0,6% | |
| Máximo | 1,30 | 4,53 | 2,71 | 9,4% | 172,0% | |
| Mínimo | -0,91 | -3,92 | -2,83 | 0,3% | -63,5% | |
| Desvio P. | 0,25 | 2,02 | 0,95 | 1,4% | 32,5% | |

Fonte: Dados de pesquisa (2021).

A matriz de correlações entre as variáveis, apresentada na Tabela 13, indica as seguintes aplicações: (i) baixa correlação entre as frequências de busca dos índices, indicando que as buscas em “Ibovespa”, “SMLL11” e “IBrX”, provavelmente, não seguem a mesma tendência de evolução ao longo do período, um dos motivos para esse fato, pode ser a concentração e a preferência de buscas do termo “Ibovespa” em relação aos demais termos; (ii) baixa correlação contemporânea entre os retornos dos índices e as frequências de busca no Google; (iii) correlação contemporânea média entre a frequência de buscas do termo “Ibovespa” e o seu volume transacionado e a sua volatilidade semanal.

TABELA 13 – Matriz de correlações – Atenção dos investidores e os indicadores de mercado

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|----------------------------|------|------|------|-------|-------|------|-------|-------|------|-------|-------|
| (1) Retorno_Ibov | 1,00 | 1,00 | 0,87 | 0,01 | -0,05 | 0,11 | 0,00 | -0,07 | 0,12 | -0,05 | 0,05 |
| (2) Retorno_IBrX | | 1,00 | 0,88 | 0,01 | -0,05 | 0,11 | -0,01 | -0,08 | 0,12 | -0,05 | 0,04 |
| (3) Retorno_SML11 | | | 1,00 | -0,03 | -0,03 | 0,07 | -0,08 | -0,06 | 0,08 | -0,10 | -0,01 |
| (4) Δ SVI_IBOV | | | | 1,00 | 0,11 | 0,05 | 0,70 | 0,14 | 0,06 | 0,25 | 0,56 |
| (5) Δ SVI_IBrX | | | | | 1,00 | 0,12 | 0,02 | 0,78 | 0,08 | -0,12 | -0,04 |
| (6) Δ SVI_SMLL11 | | | | | | 1,00 | 0,03 | 0,10 | 0,76 | -0,05 | 0,05 |
| (7) ASV_Ibov | | | | | | | 1,00 | 0,07 | 0,10 | 0,36 | 0,31 |
| (8) ASV_IBRX | | | | | | | | 1,00 | 0,08 | -0,12 | -0,01 |
| (9) ASV_SMLL | | | | | | | | | 1,00 | -0,03 | 0,04 |
| (10) VOLATILIDADE IBOVESPA | | | | | | | | | | 1,00 | 0,09 |
| (11) Δ Volume_IBOV | | | | | | | | | | | 1,00 |

Fonte: Dados de pesquisa (2021).

Por fim, conforme apresentado na Tabela 14, com base nos testes de ADF, os resultados estatísticos confirmam que cada série temporal utilizada como base de análise é estacionária, condição necessária para a eficaz utilização dos modelos de Auto Regressão Vetorial (VAR).

TABELA 14 – Teste da raiz unitária – Atenção dos investidores e os indicadores de mercado.

| Variáveis | Em nível Estatística t | Em nível p-valor | Significância de 1%: |
|-----------------------|------------------------|------------------|----------------------|
| Retorno_Ibov | -16,942 | 0.0000 | -3,462 |
| Retorno_IBrX | -16,894 | 0.0000 | -3,462 |
| Retorno_SML11 | -15,814 | 0.0000 | -3,462 |
| Δ SVI_IBOV | -14,245 | 0.0000 | -3,462 |
| Δ SVI_IBrX | -13,123 | 0.0000 | -3,462 |
| Δ SVI_SMLL11 | -11,964 | 0.0000 | -3,462 |
| ASV_Ibov | -12,029 | 0.0000 | -3,462 |
| ASV_IBRX | -11,742 | 0.0000 | -3,462 |
| ASV_SMLL | -15,558 | 0.0000 | -3,462 |
| VOLATILIDADE IBOVESPA | -13,335 | 0.0000 | -3,462 |
| Δ Volume_IBOV | -16,201 | 0.0000 | -3,462 |

Fonte: Dados de pesquisa (2021).

Complementando a análise de estacionaridade, a Figura 5 apresenta os gráficos de cada série temporal das variáveis utilizadas nessa sessão.

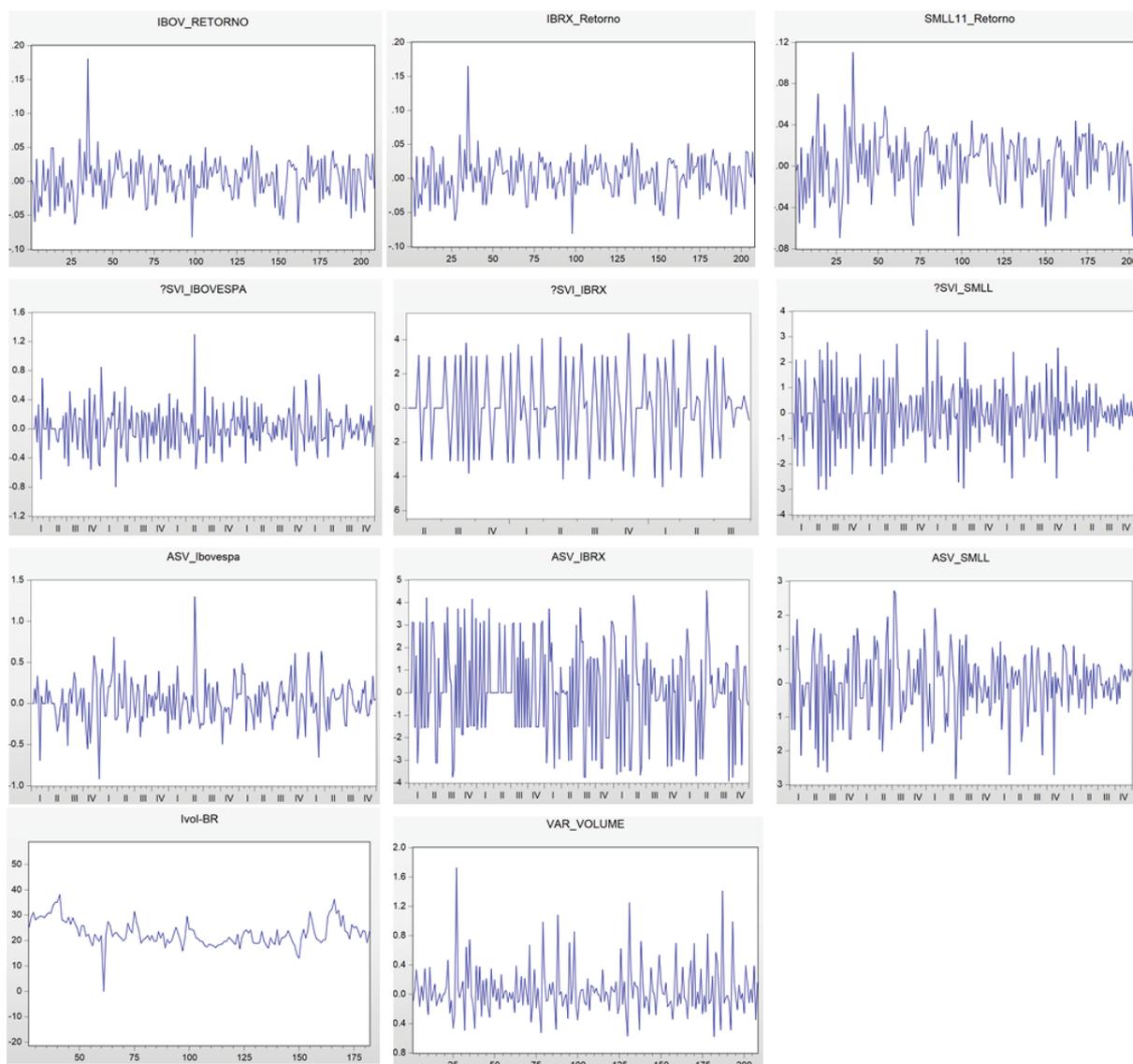


FIGURA 5 – Gráficos das séries temporais das variáveis analisadas nessa seção.
 FONTE: Autor (2021)

Assim como na seção 4, antes de executarmos as causalidades da Granger, exploramos os testes AIC, SC e HQ. Conforme valores sublinhados na Tabela 15, é possível identificar que as seleções se encontram, principalmente, dentro da especificação de até quatro períodos de defasagens. Portanto, optamos por analisar as causalidades de Granger para todos os períodos entre 1 e 4 defasagens.

TABELA 15 – Testes de defasagens – AIC, SC e HQ.

| | | Δ SVI_IBOV | | | Δ SVI_IBrX | | | Δ SVI_SMLL11 | | |
|------|---------------|-------------------|----------------|----------------|-------------------|----------------|----------------|---------------------|----------------|--|
| | | Retorno_Ibovespa | | | Retorno_IBRX-100 | | | Retorno_SML11 | | |
| LAGS | AIC | SC | HQ | AIC | SC | HQ | AIC | SC | HQ | |
| 0 | -4,04 | -4,01 | -4,03 | 0,23 | 0,26 | 0,24 | -1,33 | -1,30 | -1,32 | |
| 1 | -4,11 | -4,02 | -4,07 | -0,07 | 0,01 | -0,04 | -1,56 | -1,47 | -1,52 | |
| 2 | -4,22 | <u>-4,074*</u> | <u>-4,159*</u> | -0,22 | -0,08 | -0,16 | -1,64 | <u>-1,501*</u> | -1,59 | |
| 3 | -4,22 | -4,03 | -4,14 | -0,29 | <u>-0,087*</u> | -0,21 | -1,70 | -1,50 | <u>-1,617*</u> | |
| 4 | -4,22 | -3,96 | -4,12 | -0,32 | -0,06 | <u>-0,213*</u> | -1,70 | -1,44 | -1,60 | |
| 5 | -4,22 | -3,91 | -4,09 | -0,34 | -0,03 | -0,21 | -1,71 | -1,40 | -1,59 | |
| 6 | -4,19 | -3,83 | -4,05 | -0,32 | 0,05 | -0,17 | -1,70 | -1,33 | -1,55 | |
| 7 | -4,21 | -3,78 | -4,04 | -0,32 | 0,10 | -0,15 | -1,69 | -1,27 | -1,52 | |
| 8 | <u>-4,26*</u> | -3,78 | -4,07 | <u>-0,345*</u> | 0,13 | -0,15 | <u>-1,737*</u> | -1,26 | -1,54 | |
| 9 | -4,24 | -3,71 | -4,03 | -0,32 | 0,22 | -0,11 | -1,73 | -1,19 | -1,51 | |
| 10 | -4,22 | -3,63 | -3,98 | -0,30 | 0,29 | -0,06 | -1,72 | -1,13 | -1,48 | |
| 11 | -4,20 | -3,55 | -3,94 | -0,30 | 0,35 | -0,04 | -1,70 | -1,05 | -1,44 | |
| 12 | -4,26 | -3,55 | -3,97 | -0,28 | 0,43 | 0,00 | -1,69 | -0,99 | -1,41 | |

| | | ASV_IBOV | | | ASV_IBrX | | | ASV_SMLL11 | | |
|------|----------------|------------------|----------------|----------------|------------------|----------------|---------------|---------------|---------------|--|
| | | Retorno_Ibovespa | | | Retorno_IBRX-100 | | | Retorno_SML11 | | |
| LAGS | AIC | SC | HQ | AIC | SC | HQ | AIC | SC | HQ | |
| 0 | -4,14 | -4,11 | -4,13 | -0,09 | <u>-0,060*</u> | <u>-0,077*</u> | -1,67 | <u>-1,65*</u> | <u>-1,67*</u> | |
| 1 | -4,19 | -4,11 | -4,16 | -0,10 | -0,01 | -0,06 | <u>-1,67*</u> | -1,58 | -1,63 | |
| 2 | <u>-4,256*</u> | <u>-4,115*</u> | <u>-4,199*</u> | -0,10 | 0,04 | -0,04 | -1,66 | -1,52 | -1,60 | |
| 3 | -4,24 | -4,04 | -4,16 | -0,09 | 0,11 | -0,01 | -1,67 | -1,47 | -1,59 | |
| 4 | -4,24 | -3,98 | -4,14 | <u>-0,105*</u> | 0,15 | 0,00 | -1,66 | -1,41 | -1,56 | |
| 5 | -4,21 | -3,90 | -4,09 | -0,09 | 0,22 | 0,04 | -1,65 | -1,33 | -1,52 | |
| 6 | -4,19 | -3,82 | -4,04 | -0,06 | 0,31 | 0,09 | -1,63 | -1,26 | -1,48 | |
| 7 | -4,20 | -3,77 | -4,03 | -0,05 | 0,38 | 0,12 | -1,62 | -1,20 | -1,45 | |
| 8 | -4,24 | -3,76 | -4,05 | -0,08 | 0,41 | 0,12 | -1,65 | -1,17 | -1,46 | |
| 9 | -4,22 | -3,68 | -4,00 | -0,05 | 0,49 | 0,17 | -1,64 | -1,10 | -1,42 | |
| 10 | -4,21 | -3,61 | -3,97 | -0,02 | 0,57 | 0,22 | -1,61 | -1,02 | -1,37 | |
| 11 | -4,19 | -3,54 | -3,93 | 0,00 | 0,65 | 0,27 | -1,58 | -0,93 | -1,32 | |
| 12 | -4,24 | -3,54 | -3,96 | 0,02 | 0,73 | 0,31 | -1,59 | -0,88 | -1,30 | |

| | | VOLATILIDADE IBOVESPA | | | VOLATILIDADE IBOVESPA | | |
|------|---------------|-----------------------|---------------|---------------|-----------------------|--------------|--|
| LAGS | AIC | SC | HQ | AIC | SC | HQ | |
| 0 | 6,32 | 6,35 | 6,334 | 6,21 | 6,24 | 6,22 | |
| 1 | 5,76 | 5,85 | 5,793 | 5,64 | 5,73 | 5,68 | |
| 2 | 5,58 | <u>5,737*</u> | <u>5,646*</u> | <u>5,535*</u> | <u>5,68*</u> | <u>5,59*</u> | |
| 3 | 5,58 | 5,79 | 5,665 | 5,56 | 5,77 | 5,64 | |
| 4 | <u>5,555*</u> | 5,83 | 5,667 | 5,56 | 5,83 | 5,67 | |
| 5 | 5,57 | 5,90 | 5,702 | 5,58 | 5,92 | 5,72 | |
| 6 | 5,59 | 5,99 | 5,752 | 5,61 | 6,01 | 5,77 | |
| 7 | 5,57 | 6,03 | 5,759 | 5,60 | 6,06 | 5,79 | |
| 8 | 5,61 | 6,13 | 5,817 | 5,63 | 6,15 | 5,84 | |
| 9 | 5,62 | 6,20 | 5,854 | 5,65 | 6,23 | 5,88 | |
| 10 | 5,65 | 6,29 | 5,909 | 5,67 | 6,31 | 5,93 | |
| 11 | 5,67 | 6,37 | 5,952 | 5,68 | 6,38 | 5,96 | |
| 12 | 5,66 | 6,43 | 5,971 | 5,69 | 6,45 | 6,00 | |

| | | Δ Volume_IBOV | | | Δ Volume_IBOV | | |
|------|---------------|----------------------|----------------|---------------|----------------------|---------------|--|
| LAGS | AIC | SC | HQ | AIC | SC | HQ | |
| 0 | -5,57 | -5,54 | -5,561 | -5,75 | -5,72 | -5,74 | |
| 1 | -5,74 | -5,65 | -5,702 | -5,83 | <u>-5,74*</u> | -5,79 | |
| 2 | -5,81 | <u>-5,670*</u> | -5,755 | <u>-5,85*</u> | -5,71 | <u>-5,79*</u> | |
| 3 | <u>-5,84*</u> | -5,64 | <u>-5,759*</u> | -5,85 | -5,65 | -5,77 | |
| 4 | -5,83 | -5,58 | -5,731 | -5,83 | -5,58 | -5,73 | |
| 5 | -5,82 | -5,51 | -5,698 | -5,81 | -5,50 | -5,68 | |
| 6 | -5,82 | -5,45 | -5,670 | -5,79 | -5,43 | -5,65 | |
| 7 | -5,82 | -5,40 | -5,654 | -5,80 | -5,38 | -5,63 | |
| 8 | -5,81 | -5,33 | -5,619 | -5,79 | -5,31 | -5,60 | |
| 9 | -5,82 | -5,28 | -5,604 | -5,79 | -5,26 | -5,58 | |
| 10 | -5,80 | -5,21 | -5,564 | -5,77 | -5,17 | -5,53 | |
| 11 | -5,80 | -5,15 | -5,539 | -5,78 | -5,13 | -5,51 | |
| 12 | -5,81 | -5,10 | -5,523 | -5,79 | -5,08 | -5,50 | |

Fonte: Dados de pesquisa (2021).

Os resultados das regressões VAR e dos testes de causalidade de Granger são apresentados a seguir, divididos em 3 seções: retornos semanais, volumes semanais/anormais e volatilidade semanal.

5.1 ATENÇÃO EM ÍNDICES DE AÇÕES E O RETORNO SEMANAL

Nessa subseção, foram analisados os impactos da atenção dos investidores, medidos pela variação na frequência de busca ou por frequências anormais de busca de palavras-chave, na variação semanal dos retornos dos índices Ibovespa, IBrX-100 e SMLL11.

As regressões 5.1 e 5.2 serviram de base para os testes de causalidade de Granger entre a frequência de buscas semanal/anormal no nome dos índices de ações e o retorno semanal.

$$retorno_{i,t} = \beta_0 + \sum_{k=1}^{lags} \beta \text{atenção}_{i,t-k} + \sum_{k=1}^{lags} \alpha retorno_{i,t-k} \quad (5.1)$$

$$atenção_{i,t} = \beta_0 + \sum_{k=1}^{lags} \beta \text{atenção}_{i,t-k} + \sum_{k=1}^{lags} \alpha retorno_{i,t-k} \quad (5.2)$$

Onde, a variável atenção assume os valores do Δ SVI e ASVI e a variável retorno assume os valores do retorno semanal do Ibovespa, SMLL11 e IBrX-100.

A Tabela 16 apresenta os resultados dos testes do p-valor para cada variação semanal do retorno semanal dos índices de ações e os níveis de atenção semanais (Δ SVI) ou anormais (ASVI), conforme defasagens (de 1 a 4) utilizadas e apresentadas por coluna, em ambos os sentidos de causalidade.

TABELA 16 – Teste de causalidade de Granger – Atenção em índices e o retorno semanal

Esta tabela exibe os valores de p para testes de causalidade bilaterais da Granger sobre a atenção em índices e o retorno semanal dos índices de ações. Os resultados se referem à hipótese nula (H0) de ambas as direções relatadas no painel direito e esquerdo, respectivamente. As colunas incluem os resultados sobre as especificações de defasagem de 1 a 4 semanas. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| Lags | H0: Atenção não causa granger retorno índice | | | | H0: Retorno não causa granger atenção | | | |
|-----------------------|--|--------|--------|--------|---------------------------------------|-----------|----------|----------|
| | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| Δ SVI_IBOVESPA | 0,1806 | 0,3477 | 0,2362 | 0,3304 | 0,4107 | 0,0836* | 0,2011 | 0,1642 |
| ASVI_IBOVESPA | 0,943 | 0,2425 | 0,3760 | 0,3090 | 0,5082 | 0,0533* | 0,0978* | 0,0481** |
| Δ SVI_SMLL | 0,7698 | 0,8742 | 0,9663 | 0,9913 | 0,2110 | 0,3260 | 0,1802 | 0,1546 |
| ASV_SMLL | 0,6918 | 0,8521 | 0,9495 | 0,9519 | 0,1038 | 0,3240 | 0,1134 | 0,2050 |
| Δ SVI_IBRX | 0,2778 | 0,2072 | 0,3050 | 0,2782 | 0,0027*** | 0,0207** | 0,0122** | 0,0376** |
| ASVI_IBRX | 0,9622 | 0,3157 | 0,4531 | 0,6341 | 0,0048*** | 0,0062*** | 0,008*** | 0,0183** |

Fonte: Dados de pesquisa (2021).

Conforme tabela acima, não foram identificados resultados significativos para a análise de causalidade entre as variáveis Δ SVI (evolução das buscas em relação à semana anterior) e ASV (evolução anormal das buscas no Google) e o retorno semanal dos principais índices de mercado. No sentido oposto de causalidade, os retornos dos índices Ibovespa e IBrX-100 causaram Granger nas variáveis Δ SVI e ASV, indicando que uma variação no retorno é seguida de um maior ou menor interesse dos investidores nesses índices.

Com o objetivo de complementar os resultados identificados no teste de causalidade de Granger, para a equação 5.1, realizamos estimativas do modelo VAR com defasagens ótimas. Os coeficientes seguidos pelos erros padrões são apresentados na Tabela 17.

TABELA 17 – Estimativa VAR - Atenção em índices e o retorno semanal

Esta tabela fornece resultados de estimativa VAR para os índices de ações sobre valores defasados da atenção dos investidores e a variação do retorno semanal dos índices. As duas especificações do VAR incluem defasamentos conforme testes de AR e LC. Os coeficientes reportados são seguidos por erros padrão entre parênteses. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| Δ SVI | | | | ASVI | | | |
|------------------|-------------------------|-------------------------|-------------------------|-------------|-------------------------|-------------------------|-------------------------|
| | Ret Ibov | Ret IBRX | Ret Sml11 | | Ret Ibov | Ret IBRX100 | Ret Sml11 |
| Variáveis | Max Lag = 3 | Max lag =4 | Max lag =3 | Variáveis | Max Lag = 2 | Max lag =2 | Max lag =3 |
| Δ SVI t-1 | -0,005599 (0,007456) | -0,000321 (0,00099) | -0,000671 (0,001776) | ASVI t-1 | -0,004131 (0,007327) | 0,000031 (0,000899) | -0,00102 (0,001779) |
| Δ SVI t-2 | 0,006329 (0,007667) | 0,001294 (0,001225) | -0,000495 (0,002016) | ASVI t-2 | 0,012318* (0,007321) | 0,001279 (0,000888) | -0,000247 (0,001769) |
| Δ SVI t-3 | 0,010774 (0,007441) | 0,000482 (0,001225) | 0,000269 (0,001765) | ASVI t-3 | | -0,000486 (0,000891) | 0,000004 (0,001764) |
| Δ SVI t-4 | | 0,001386 (0,00097) | | ASVI t-4 | | 0,000426 (0,000882) | |
| Retorno t-1 | -0,048062 (0,062845) | -0,034786 (0,063194) | 0,011024 (0,063164) | Retorno t-1 | -0,047306 (0,062058) | -0,042615 (0,063310) | 0,011625 (0,063182) |
| Retorno t-2 | 0,115982* (0,062567) | 0,106877* (0,063967) | 0,101591 (0,063136) | Retorno t-2 | 0,106997* (0,062119) | 0,110067 (0,064282) | 0,10314 (0,063216) |
| Retorno t-3 | 0,01392 (0,063358) | 0,002582 (0,064103) | 0,06234 (0,063355) | Retorno t-3 | | -0,000913 (0,064406) | 0,065022 (0,063418) |
| Retorno t-4 | | -0,020827 (0,064178) | | Retorno t-4 | | -0,023502 (0,064517) | |
| Constante | 0,00341 (0,001864) | 0,003785 (0,001806) | 0,003752 (0,001727) | Constante | 0,003356 (0,001851) | 0,003764 (0,001820) | 0,0037 (0,0017) |
| Obs. | 258 | 257 | 258 | Obs. | 259 | 257 | 258 |
| R2 ajust. | 0,007366 | 0,004092 | -0,008063 | R2 ajust. | 0,024801 | -0,006083 | -0,007708 |

Fonte: Dados de pesquisa (2021).

A relação entre a atenção anormal (ASVI) em “Ibovespa” e o retorno semanal do índice (Ret_Ibov) indicou que a variação na frequência de buscas possui um impacto positivo no retorno semanal dos índices de ação Ibovespa de duas semanas posteriores, com significância de 10%. Isto sugere que as pesquisas são utilizadas antes de uma decisão de compra por parte dos potenciais acionistas, o que, por sua vez, elevará os preços das ações pertencentes ao índice. Esse comportamento estatisticamente significativo se mostra alinhado a hipótese “price pressure hypothesis”, de Barber e Odean (2008).

Entretanto, exceto pela relação entre o retorno do Ibovespa e atenção anormal (ASVI), os resultados obtidos indicaram, que os termos de buscas do nome dos índices de ações no Google Trends, não se mostraram uma ferramenta eficiente de interesse de investidores. Essa ausência de significância estatística é equivalente a

identificada por Caraça (2019), ao analisar a relação da frequência semanal de buscas (Δ SVI) e o termo “Ibovespa” e o retorno do índice.

Dessa forma, mesmo que a relação tenha significância estatística apenas para uma das estimativas realizadas, os resultados suportam a hipótese H2a, definida como “um aumento na atenção dos investidores nos principais índices do mercado brasileiro reflete uma mudança positiva ou negativo no retorno dos índices”.

5.2 ATENÇÃO EM ÍNDICES DE AÇÕES E A VOLATILIDADE SEMANAL

Nessa subseção analisamos a relação entre frequência de buscas semanal/anormal no nome dos índices de ações e a volatilidade semanal. As equações base para os testes de causalidade de Granger seguiram a metodologia aplicada nas seções 4.1, 4.2, 4.3 e 5.1, sendo representada pelas equações 5.3 e 5.4

$$volatilidade_{i,t} = \beta_0 + \sum_{k=1}^{lags} \beta \text{atenção}_{i,t-k} + \sum_{k=1}^{lags} \alpha volatilidade_{i,t-k} \quad (5.3)$$

$$\text{atenção}_{i,t} = \beta_0 + \sum_{k=1}^{lags} \beta \text{atenção}_{i,t-k} + \sum_{k=1}^{lags} \alpha volatilidade_{i,t-k} \quad (5.4)$$

Onde, a variável atenção assume os valores do Δ SVI e ASVI do termo “Ibovespa” e a variável volatilidade assume os valores da volatilidade semanal do Ibovespa. Para os demais índices de mercado, não foi possível analisar essa variação, devido a limitação de acesso aos conjuntos de dados.

A Tabela 18 indica os resultados dos testes do p-valor para cada variação semanal da volatilidade semanal do Ibovespa e os níveis de atenção, conforme defasagens (de 1 a 4) utilizadas e apresentadas por coluna, em ambos os sentidos de causalidade.

TABELA 18 – Teste de causalidade de Granger – Atenção em índices e a volatilidade semanal
Esta tabela exibe os valores de p para testes de causalidade bilaterais da Granger sobre a atenção em índices e a volatilidade semanal dos índices de ações. Os resultados se referem à hipótese nula (H0) de ambas as direções relatadas no painel direito e esquerdo, respectivamente. As colunas incluem os resultados sobre as especificações de defasagem de 1 a 4 semanas. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| Lags | H0: Atenção não causa granger volatilidade índice | | | | H0: Volatilidade não causa granger Atenção | | | |
|-----------------------|---|--------|--------|--------|--|--------|--------|--------|
| | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| Δ SVI_IBOVESPA | 0.4586 | 0.5283 | 0.5969 | 0.7455 | 0.0132** | 0.2275 | 0.5875 | 0.4787 |
| ASVI_IBOVESPA | 0.2786 | 0.4853 | 0.6374 | 0.5463 | 0.5944 | 0.6938 | 0.7204 | 0.7197 |

Fonte: Dados de pesquisa (2021).

Os resultados da causalidade de Granger confirmaram somente a hipótese de que a volatilidade causa Granger da variação semanal da frequência de buscas (atenção Δ SVI), com significância de 5% e defasagem de uma semana. Isso significa que variações na volatilidade semanal do índice Ibovespa refletem na semana seguinte uma variação na frequência de buscas da palavra “Ibovespa”. No sentido contrário dessa relação, nenhum p-valor significativamente relevante foi identificado.

Adicionalmente, para a equação 5.3 realizamos estimativas do modelo VAR com defasagens ótimas. Os coeficientes de regressão e os erros padrões são expressos na Tabela 19.

TABELA 19 – Estimativa VAR - Atenção em índices e a volatilidade semanal

Esta tabela fornece resultados de estimativa VAR para os índices de ações sobre valores defasados da atenção dos investidores e a volatilidade semanal dos índices. As duas especificações do VAR incluem defasamentos conforme testes de AR e LC. Os coeficientes reportados são seguidos por erros padrão entre parênteses. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| | Δ SVI | | ASVI |
|------------------|--------------------------|------------------|--------------------------|
| | Vol. Ibovespa | | Vol. Ibovespa |
| Variáveis | Max Lag = 3 | Variáveis | Max Lag = 2 |
| Δ SVI t-1 | 0,004587 (0,00366) | ASVI t-1 | 0,004316 (0,00366) |
| Δ SVI t-2 | 0,001760 (0,003776) | ASVI t-2 | -0,001918 (0,003665) |
| Δ SVI t-3 | 0,003099 (0,003514) | ASVI t-3 | |
| Volatilidade t-1 | 0,143517** (0,066582) | Volatilidade t-1 | 0,149593** (0,067123) |
| Volatilidade t-2 | 0,039783 (0,067576) | Volatilidade t-2 | 0,070224 (0,066754) |
| Volatilidade t-3 | 0,115673* (0,06585) | Volatilidade t-3 | |
| Constante | 0,019541 (0,00277) | Constante | 0,021875 (0,002516) |
| Obs. | 258 | Obs. | 259 |
| R2 ajust. | 0,035673 | R2 ajust. | 0,028624 |

Fonte: Dados de pesquisa (2021).

Confirmando os resultados identificados nos testes de causalidade de Granger entre atenção dos investidores (Δ SVI e ASVI) e a volatilidade do Ibovespa, não foi identificada nenhuma relação significativa entre as variáveis analisadas. Portanto, é possível concluir que a utilização de dados do Google Trends, no que tange a palavra “Ibovespa”, não auxilia na previsibilidade da volatilidade de mercado.

Essa ausência de significância estatística é equivalente a identificada por Caraça (2019), ao analisar a relação da frequência semanal de buscas (Δ SVI) e o termo “Ibovespa” e a volatilidade do mercado.

Dessa forma, com essa base de análise, não é possível confirmar a hipótese H2b, de que a atenção dos investidores reflete uma variação na volatilidade dos índices.

5.3 ATENÇÃO EM ÍNDICES DE AÇÕES E O VOLUME SEMANAL/ANORMAL

Dedica-se essa subseção ao estudo da relação entre os volumes semanais (Δ volume) e volumes anormais (Abn_Vol) do Ibovespa e a atenção semanal ou anormal dos investidores no termo “Ibovespa”. Conforme metodologia aplicada na seção 4.1, 4.2, 4.3, 5.1 e 5.2 as regressões 5.5 e 5.6 serviram de base para os testes de causalidade de Granger.

$$volume_{i,t} = \beta_0 + \sum_{k=1}^{lags} \beta \text{atenção}_{i,t-k} + \sum_{k=1}^{lags} \alpha volume_{i,t-k} \quad (5.5)$$

$$\text{atenção}_{i,t} = \beta_0 + \sum_{k=1}^{lags} \beta \text{atenção}_{i,t-k} + \sum_{k=1}^{lags} \alpha volume_{i,t-k} \quad (5.6)$$

Onde, a variável atenção assume os valores do Δ SVI ou ASVI e a variável volume assume os valores do volume semanal ou volume anormal do Ibovespa. Para os demais índices de mercado, não foi possível analisar essa variação, devido a limitação de acesso aos conjuntos de dados.

A Tabela 20 indica os resultados dos testes do p-valor para cada variação semanal do volume semanal ou anormal do Ibovespa e os níveis de atenção, conforme defasagens (de 1 a 4) utilizadas e apresentadas por coluna, em ambos os sentidos de causalidade.

TABELA 20 – Teste de causalidade de Granger – Atenção em índices e o volume semanal e anormal. Esta tabela exibe os valores de p para testes de causalidade bilaterais da Granger sobre a atenção em índices e o volume semanal dos índices de ações. Os resultados se referem à hipótese nula (H0) de ambas as direções relatadas no painel direito e esquerdo, respectivamente. As colunas incluem os resultados sobre as especificações de defasagem de 1 a 4 semanas. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| H0: Atenção não causa granger volume anormal | | | | | H0: Volume anormal não causa granger atenção | | | |
|--|--------|---------|--------|--------|--|--------|--------|--------|
| Lags | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| Δ SVI_IBOVESPA | 0.4345 | 0.1959 | 0.2774 | 0.2726 | 0.9262 | 0.4734 | 0.5114 | 0.4542 |
| ASVI_IBOVESPA | 0.5288 | 0.0939* | 0.1655 | 0.2242 | 0.0436** | 0.2331 | 0.4278 | 0.5115 |

| H0: Atenção não causa granger volume semanal | | | | | H0: Volume semanal não causa granger atenção | | | |
|--|------------|------------|-----------|-----------|--|----------|-----------|----------|
| Lags | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| Δ SVI_IBOVESPA | 0.2862 | 0.0112** | 0.0179** | 0.0482** | 0.0764** | 0.0394** | 0.0053*** | 0.0235** |
| ASVI_IBOVESPA | 0,00001*** | 0,00001*** | 0.0005*** | 0.0072*** | 0.0386** | 0.2154 | 0.0349** | 0.1790 |

Fonte: Dados de pesquisa (2021).

Analisando os resultados identificados na tabela acima, primeiramente, para a variável que mede a variação anormal do volume negociado do índice Ibovespa, identificamos que ela causa Granger na variável frequência de buscas anormais da palavra Ibovespa com uma semana de defasagem, com significância de 5%. Isso significa que oscilações de volume fora do padrão de negociação geram uma mudança no comportamento na frequência de buscas no Google. Na relação inversa, identificamos que para uma defasagem de duas semanas, a frequência anormal de buscas causa Granger do volume anormal, com representatividade de 10%.

Já para a variável variação do volume semanal, identificamos uma forte relação com as variáveis de atenção, em ambos os sentidos de causalidade. Por exemplo, com nível de significância de 5%, verificamos que o Δ SVI causa Granger do variável volume semanal, ou seja, um aumento no nível de buscas no Google prediz um movimento no volume negociado.

Adicionalmente, para a equação 5.5 realizamos estimativas do modelo VAR com defasagens ótimas. Os coeficientes seguidos pelos erros padrões são apresentados na Tabela 21.

TABELA 21 – Estimativa VAR - Atenção em índices e o volume semanal e anormal

Esta tabela fornece resultados de estimativa VAR para os índices de ações sobre valores defasados da atenção dos investidores e o volume semanal dos índices. As duas especificações do VAR incluem defasamentos conforme testes de AR e LC. Os coeficientes reportados são seguidos por erros padrão entre parênteses. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| Abn_v Ibov | | Abn_v Ibov | | Volume Ibov | | Volume Ibov | |
|------------------|----------------------------|------------|----------------------------|---------------------|---------------------------|---------------------|--------------------------|
| Variáveis | Max Lag = 2 | Variáveis | Max Lag = 2 | Variáveis | Max Lag = 3 | Variáveis | Max Lag = 3 |
| Δ SVI t-1 | -0,292043* (0,165407) | ASVI t-1 | -0.363507** (0.172578) | Δ SVI t-1 | -0,173578* (0,091026) | ASVI t-1 | -0,217656** 0,087779 |
| Δ SVI t-2 | -0,150715 (0,164521) | ASVI t-2 | -0.000174 (0.173944) | Δ SVI t-2 | -0,300321*** (0,09628) | ASVI t-2 | -0,220396** 0,092797 |
| Δ SVI t-3 | | ASVI t-3 | | Δ SVI t-3 | -0,147806 (0,092339) | ASVI t-3 | 0,033576 0,086785 |
| Δ SVI t-4 | | ASVI t-4 | | Δ SVI t-4 | | ASVI t-4 | |
| Abn_V t-1 | 0,845252*** (0,059663) | Abn_V t-1 | 0.842236*** (0.059668) | Δ volume t-1 | -0,355635*** (0,07414) | Δ volume t-1 | -0,368336*** 0,072761 |
| Abn_V t-2 | -0,338949*** (0,059516) | Abn_V t-2 | -0.332654*** (0.059627) | Δ volume t-2 | -0,138026* (0,07957) | Δ volume t-2 | -0,143641** 0,076949 |
| Abn_V t-3 | | Abn_V t-3 | | Δ volume t-3 | -0,071281 (0,075194) | Δ volume t-3 | -0,07157 0,064665 |
| Abn_V t-4 | | Abn_V t-4 | | Δ volume t-4 | | Δ volume t-4 | |
| Constante | -0,053134 (0,04277) | Constante | -0.047433 (0,04277) | Constante | 0,076637 (0,019594) | Constante | 0,081949 0,019524 |
| Obs. | 259 | Obs. | 259 | Obs. | 258 | Obs. | 258 |
| R2 ajust. | 0,453794 | R2 ajust. | 0.456946 | R2 ajust. | 0,176287 | R2 ajust. | 0,200346 |

Fonte: Dados de pesquisa (2021).

Conforme resultados observados acima, nas quatro estimativas realizadas, identificamos relações significativas entre as variáveis de atenção defasadas e os volumes semanais ou anormais de negociação do mercado, sendo as principais conclusões: (i) um menor nível de buscas semanais e anormais do índice Ibovespa, com uma semana de defasagem, prediz um aumento no volume de negociações anormais do Ibovespa; (ii) menores níveis de buscas semanais e anormais do índice Ibovespa, com uma e duas semanas de defasagem, predizem um aumento no volume de negociações semanais do Ibovespa.

Isso indica que os investidores buscam menos informações sobre o índice Ibovespa antes de realizarem operações de compra e venda no mercado, confirmando a hipótese H2c, de que a atenção dos investidores reflete uma variação no volume negociado no mercado.

6 ATENÇÃO EM TICKERS/EMPRESAS E OS INDICADORES DE MERCADO

Nessa seção investigamos se os retornos anormais, retornos semanais, volumes anormais e volumes semanais de negociação podem ser previstos pela variação da frequência de buscas no Google (anormal ou semanal) dos *tickers* de negociação dos títulos, dos nomes das empresas selecionadas e dos nomes das empresas seguidos da palavra “ação”.

Para as análises utilizamos regressões com dados em painel, formados por 74 ações, selecionadas conforme seção 3.1.1, ao longo de 261 semanas (5 anos), totalizando 19.314 observações. É importante ressaltar que as estimativas incluíram as defasagens da evolução das buscas em relação à semana anterior (Δ SVI) ou a evolução das buscas em relação a mediana de 4 semanas (ASVI) como variáveis explicativas, conforme metodologia aplicada nos trabalhos de Bijl et al. (2016) e Takeda e Wakao (2014).

Adicionalmente, optamos pela realização do teste de Hausman para verificar se os modelos se adequam melhor ao modelo de efeitos fixos ou modelo de efeitos aleatórios. Caso se aceite a hipótese H_0 do teste de Hausman, o modelo de efeitos aleatórios é o mais adequado a ser empregado, caso se rejeite a hipótese H_0 , o modelo de efeitos fixo se torna mais aderente. No presente trabalho, os resultados indicaram que as regressões com dados em painéis se adaptaram aos efeitos fixos ao rejeitarem a hipótese nula.

A Tabela 16 apresenta as principais estatísticas descritivas para as variáveis das regressões. Destacando que o retorno semanal médio das ações foi em torno de 0,6%, evidenciando que a seleção de dados apresentou viés de alta no período analisado, e que a mediana das variações das frequências de buscas foi zero, indicando que determinados termos buscados (*ticker*, nome da empresa ou nome da empresa + “ações”) tiveram períodos sem busca no Google, como por exemplo, o *ticker* “ITSA4” no início de 2015. Outro ponto de destaque são os pontos de máxima e mínima de retorno semanal e volume semanal, podendo ser explicados por eventuais movimentações atípicas de compra e venda de determinado ativo.

TABELA 22 – Estatísticas descritivas – Dados em painéis

| | RET | ABN RET | ABN VOL | Δ Vol | Δ SVI ACOES | Δ SVI EMPRESA | Δ SVI TICKER | ASVI ACOES | ASVI EMPRESA | ASVI_ TICKER |
|--------|--------|------------|------------|--------------|-----------------------|-------------------------|------------------------|---------------|-----------------|-----------------|
| Média | 0,6% | 0,4% | -2% | 17% | 0,008 | 0,000 | 0,010 | 0,003 | -0,001 | 0,003 |
| Median | 0,4% | 0,4% | -1% | -1% | 0,000 | 0,000 | 0,000 | 0,000 | 0,000 | 0,000 |
| Máximo | 97,5% | 3,5% | 865% | 53400% | 4,605 | 4,605 | 4,094 | 6,561 | 6,056 | 5,699 |
| Mínimo | -54,5% | -4,1% | -1812% | -100% | -4,605 | -4,605 | -4,431 | -6,871 | -6,893 | -5,102 |
| Desv_P | 5,6% | 0,7% | 55% | 407% | 1,565 | 0,687 | 0,764 | 1,781 | 0,785 | 0,876 |

Fonte: Dados de pesquisa (2021).

A matriz de correlações das variáveis, apresentada na Tabela 23, indicou, principalmente, que a correlação contemporânea entre os indicadores de frequência do Google (Δ SVI e ASVI) e os retornos (semanais e anormais) é baixa, bem como, a correlação entre os indicadores calculados para cada uma das categorias dos termos buscados (*ticker*, nome da empresa ou nome da empresa + “ações”). Por outro lado, como esperado, a correlação entre os indicadores de evolução de buscas semanais e anormais é forte para todos os grupos de palavras analisados.

TABELA 23 – Matriz de Correlação – Dados em painéis

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|--------------------------|-------|-------|-------|-------|--------|--------|-------|--------------|--------------|--------------|
| (1) RETORNO_SEMANAL | 1,000 | 0,015 | 0,020 | 0,033 | -0,009 | 0,117 | 0,050 | 0,015 | 0,022 | 0,026 |
| (2) Δ SVI_ACOES | | 1,000 | 0,053 | 0,043 | 0,013 | 0,001 | 0,047 | 0,968 | 0,051 | 0,041 |
| (3) Δ SVI_EMPRESA | | | 1,000 | 0,044 | 0,012 | 0,001 | 0,045 | 0,051 | 0,966 | 0,043 |
| (4) Δ SVI_TICKER | | | | 1,000 | 0,033 | -0,003 | 0,116 | 0,042 | 0,042 | 0,962 |
| (5) Δ Volume | | | | | 1,000 | -0,008 | 0,151 | 0,012 | 0,011 | 0,031 |
| (6) ABN_RET | | | | | | 1,000 | 0,001 | -0,001 | -0,002 | -0,001 |
| (7) ABN_VOLUME | | | | | | | 1,000 | 0,040 | 0,036 | 0,096 |
| (8) ASVI_ACOES | | | | | | | | 1,000 | 0,051 | 0,043 |
| (9) ASVI_EMPRESA | | | | | | | | | 1,000 | 0,044 |
| (10) ASVI_TICKER | | | | | | | | | | 1,000 |

Fonte: Dados de pesquisa (2021).

Os resultados das regressões são apresentados a seguir, divididos em 4 seções: retornos anormais, retornos semanais, volumes anormais e volumes semanais. Em razão da limitação dos dados, optou-se por não realizar análise da volatilidade das ações.

6.1 ATENÇÃO EM TICKERS/EMPRESAS E OS RETORNOS ANORMAIS

Dedica-se essa subseção ao estudo das relações entre a atenção semanal ou anormal dos investidores nos *tickers* de negociação, nome das empresas e nome da empresa + ações e os retornos anormais dos títulos de ações, calculado com base em uma janela de 52 semanas.

As variáveis explicativas selecionadas na regressão com dados em painéis do retorno anormal (*Abn_retorno*) foram: uma defasagem da variável de retorno semanal (*Retorno*), cinco defasagens de ΔSVI ou *ASVI* e uma defasagem da variação do volume semanal ($\Delta volume$). O modelo é, portanto, especificado conforme equações 6.1 e 6.2, sendo que a sua formação varia de acordo com a palavra utilizada para medir a atenção dos investidores (*tickers*, nome das empresas e nome das empresas + "ações".)

$$Abn_retorno_t = \alpha + \beta_1 Retorno_{t-1} + \sum_{n=1}^5 \beta_2 \Delta SVI_{t-n} + \beta_3 \Delta volume_{t-1} \quad (6.1)$$

$$Abn_retorno_t = \alpha + \beta_1 Retorno_{t-1} + \sum_{n=1}^5 \beta_2 ASVI_{t-n} + \beta_3 \Delta volume_{t-1} \quad (6.2)$$

Os resultados são apresentados na Tabela 24, divididos conforme o medidor de atenção dos investidores (semanal ou anormal) e a categoria do termo buscado no Google.

TABELA 24 – Dados em painéis - Atenção em empresas/tickers e retornos anormais

A tabela fornece os resultados das regressões com dados em painéis para o retorno anormal de um conjunto de ações sobre valores defasados do retorno semanal da ação (retorno), da frequência de buscas no Google (Δ SVI ou ASVI) e do volume semanal de negociação da ação (Δ volume). As regressões foram estimadas para as categorias de pesquisa ticker, nome das empresas e nome das empresas + "ações". Os coeficientes reportados são seguidos por erros padrão entre parênteses. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| Ticker | | | | Empresa | | | |
|---------------------|----------------------------|---------------------|---------------------------|---------------------|---------------------------|---------------------|---------------------------|
| Abn retorno | | Abn retorno | | Abn retorno | | Abn retorno | |
| Retorno t-1 | 0.012520*** (1.20E-05) | Retorno t-1 | 0.012474*** (0.000889) | Retorno t-1 | 0.012448*** (0.000888) | Retorno t-1 | 0.012441*** (0.000888) |
| Δ SVI t-1 | -0.000148* (7.99E-05) | ASVI t-1 | -5.50E-05 (6.78E-05) | Δ SVI t-1 | 5.98E-05 (8.90E-05) | ASVI t-1 | -5.09E-05 (7.46E-05) |
| Δ SVI t-2 | -0.000216** (9.72E-05) | ASVI t-2 | -6.23E-05 (7.89E-05) | Δ SVI t-2 | 0.000103 (0.000109) | ASVI t-2 | -6.00E-05 (8.66E-05) |
| Δ SVI t-3 | -0.000258** (0.000103) | ASVI t-3 | -7.62E-05 (8.19E-05) | Δ SVI t-3 | 7.76E-05 (0.000115) | ASVI t-3 | -9.28E-05 (8.96E-05) |
| Δ SVI t-4 | -0.000230** (9.63E-05) | ASVI t-4 | -7.03E-05 (7.83E-05) | Δ SVI t-4 | 5.94E-05 (0.000109) | ASVI t-4 | -9.60E-05 (8.67E-05) |
| Δ SVI t-5 | -0.000158** (7.87E-05) | ASVI t-5 | -5.80E-05 (6.70E-05) | Δ SVI t-5 | 2.64E-05 (8.92E-05) | ASVI t-5 | -8.17E-05 (7.48E-05) |
| Δ volume t-1 | 0.00000236** (0.000889) | Δ volume t-1 | -2.41E-05** (1.20E-05) | Δ volume t-1 | -2.44E-05** (1.20E-05) | Δ volume t-1 | -2.43E-05** (1.20E-05) |
| Const | 0.003774** (4.94E-05) | Const | 0.003764*** (4.93E-05) | Const | 0.003763*** (4.93E-05) | Const | 0.003763*** (4.93E-05) |
| Obs. | 18944 | Obs. | 18944 | Obs. | 18944 | Obs. | 18944 |
| Periodos | 256 | Periodos | 256 | Periodos | 256 | Periodos | 256 |
| R2 ajust. | 0.155592 | R2 ajust. | 0.155266 | R2 ajust. | 0.155242 | R2 ajust. | 0.155282 |

| Empresa + "ações" | | | |
|---------------------|---------------------------|---------------------|---------------------------|
| Abn retorno | | Abn retorno | |
| Retorno t-1 | 0.012450*** (0.000888) | Retorno t-1 | 0.012493*** (0.000888) |
| Δ SVI t-1 | -1.81E-06 (3.99E-05) | ASVI t-1 | -5.36E-05 (3.45E-05) |
| Δ SVI t-2 | 4.16E-06 (5.01E-05) | ASVI t-2 | -7.59E-05* (4.18E-05) |
| Δ SVI t-3 | 9.19E-07 (5.28E-05) | ASVI t-3 | -8.71E-05** (4.37E-05) |
| Δ SVI t-4 | 4.58E-06 (5.02E-05) | ASVI t-4 | -7.61E-05* (4.20E-05) |
| Δ SVI t-5 | -1.71E-06 (4.02E-05) | ASVI t-5 | -5.51E-05 (3.47E-05) |
| Δ volume t-1 | -2.43E-05** (1.20E-05) | Δ volume t-1 | -2.41E-05** (1.20E-05) |
| Constante | 0.003763*** (4.93E-05) | Const | 0.003764*** (4.93E-05) |
| Obs. | 18944 | Obs. | 18944 |
| Periodos | 256 | Periodos | 256 |
| R2 ajust. | 0.155203 | R2 ajust. | 0.155449 |

Fonte: Dados de pesquisa (2021).

Os resultados das estimativas indicam que, por exemplo, o retorno anormal das ações é influenciado negativamente pela variação na frequência semanal de buscas do *ticker* dos títulos no Google, medidos pela variável Δ SVI, nas 5 semanas anteriores, com significâncias aos níveis de 5% e 10%. Ou seja,

Nesse caso, o somatório dos coeficientes defasados da variável Δ SVI foi de 0,0011, significando que um aumento de 1 p.p na atenção semanal é seguido de uma queda no retorno anormal de 0,011 p.p. Apesar do pequeno impacto, os resultados indicam que os investidores procuram informações das empresas antes de efetivar uma venda no mercado.

Outra relação que teve significância estatística foi entre a variação anormal da frequência de buscas do nome da empresa seguido da palavra ação, medidas pela variável ASVI, e o retorno anormal. Os resultados indicaram uma influência negativa entre as variáveis, com significância ao nível de 10%.

Esses comportamentos estatisticamente significativos se mostraram contrários a hipótese “*price pressure hypothesis*”, de Barber e Odean (2008), e aos resultados de Da et al. (2011), que sugerem que uma maior atenção dos investidores resultaria em uma pressão positiva no preço dos títulos, ou seja, para se confirmar essa hipótese, os sinais das variáveis de atenção deveriam ser opostos aos identificados. Por outro lado, os estudos de Bijl et al. (2016), para mercado norte americano, e Yoshinaga e Rocco (2019), para o mercado brasileiro, encontraram relações similares as identificadas no presente trabalho.

As demais regressões não indicaram significância na relação entre Google Trends e retornos anormais.

Os resultados identificados, no caso dos retornos, suportam a hipótese H3a, mas rejeitam a hipótese H4a, uma vez que atenção no nome da empresa não refletiu mudanças significativas nos retornos anormais dos títulos.

6.2 ATENÇÃO EM TICKERS/EMPRESAS E OS RETORNOS SEMANAIS

Seguindo a mesma metodologia aplicada na sessão 6.1, nesse tópico investigamos a relação entre a atenção e o retorno semanal dos títulos. As regressões da variável dependente retorno semanal (retorno) são especificadas nas equações 6.3 e 6.4, onde as variáveis explicativas de atenção assumem valores defasados conforme parâmetros de pesquisa (ticker, nome das empresas ou nome das empresas + “ações”).

$$\text{Retorno}_t = \alpha + \beta_1 \text{Retorno}_{t-1} + \sum_{n=1}^5 \beta_2 \Delta \text{SVI}_{t-n} + \beta_3 \Delta \text{volume}_{t-1} \quad (6.3)$$

$$\text{Retorno}_t = \alpha + \beta_1 \text{Retorno}_{t-1} + \sum_{n=1}^5 \beta_2 \text{ASVI}_{t-n} + \beta_3 \Delta \text{volume}_{t-1} \quad (6.4)$$

Os resultados são apresentados na Tabela 25, divididos conforme o parâmetro de atenção dos investidores (semanal ou anormal) e a categoria do termo buscado no Google.

TABELA 25 – Dados em painéis - Atenção em empresas/tickers e retornos semanais

A tabela fornece os resultados das regressões com dados em painéis para o retorno semanal de um conjunto de ações sobre valores defasados do retorno semanal da ação (retorno), da frequência de buscas no Google (Δ SVI ou ASVI) e do volume semanal de negociação da ação (Δ volume). As regressões foram estimadas para as categorias de pesquisa ticker, nome das empresas e nome das empresas + "ações". Os coeficientes reportados são seguidos por erros padrão entre parênteses. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| Ticker | | Empresa | |
|---------------------|----------------------------|---------------------|----------------------------|
| Retorno semanal | Retorno semanal | Retorno semanal | Retorno semanal |
| Retorno t-1 | 0.005585 (0.007275) | Retorno t-1 | 0.005575 (0.007273) |
| Δ SVI t-1 | -0.000111 (0.000654) | ASVI t-1 | -0.000111 (0.000555) |
| Ticker | (0.000654) | Ticker | (0.000555) |
| Δ SVI t-2 | 0.000369 (0.000796) | ASVI t-2 | 0.000353 (0.000646) |
| Ticker | (0.000796) | Ticker | (0.000646) |
| Δ SVI t-3 | -5.00E-07 (0.000840) | ASVI t-3 | 0.000180 (0.000670) |
| Ticker | (0.000840) | Ticker | (0.000670) |
| Δ SVI t-4 | -0.000164 (0.000788) | ASVI t-4 | 8.78E-05 (0.000641) |
| Ticker | (0.000788) | Ticker | (0.000641) |
| Δ SVI t-5 | -0.000432 (0.000644) | ASVI t-5 | -0.000126 (0.000549) |
| Ticker | (0.000644) | Ticker | (0.000549) |
| Δ volume t-1 | -0.000268*** (9.81E-05) | Δ volume t-1 | -0.000268*** (9.81E-05) |
| Const | 0.005737*** (0.000405) | Const | 0.005732*** (0.000404) |
| Obs. | 18944 | Obs. | 18944 |
| Periodos | 256 | Periodos | 256 |
| R2 ajust. | 0 | R2 ajust. | 0 |
| Retorno t-1 | 0.005380 (0.007268) | Retorno t-1 | 0.005228 (0.007268) |
| Δ SVI t-1 | -0.001747** (0.000728) | ASVI t-1 | -0.000982 (0.000610) |
| Empresa | (0.000728) | Empresa | (0.000610) |
| Δ SVI t-2 | -0.001783** (0.000894) | ASVI t-2 | -0.001083 (0.000709) |
| Empresa | (0.000894) | Empresa | (0.000709) |
| Δ SVI t-3 | -0.002655*** (0.000938) | ASVI t-3 | -0.001919*** (0.000733) |
| Empresa | (0.000938) | Empresa | (0.000733) |
| Δ SVI t-4 | -0.003871*** (0.000893) | ASVI t-4 | -0.003211*** (0.000709) |
| Empresa | (0.000893) | Empresa | (0.000709) |
| Δ SVI t-5 | -0.002258*** (0.000730) | ASVI t-5 | -0.002172*** (0.000612) |
| Empresa | (0.000730) | Empresa | (0.000612) |
| Δ volume t-1 | -0.000267*** (9.80E-05) | Δ volume t-1 | -0.000268*** (9.80E-05) |
| Const | 0.005750*** (0.000403) | Const | 0.005738*** (0.000403) |
| Obs. | 18944 | Obs. | 18944 |
| Periodos | 256 | Periodos | 256 |
| R2 ajust. | 0.000113 | R2 ajust. | 0.000189 |

| Empresa + "ações" | |
|---------------------|----------------------------|
| Retorno semanal | Retorno semanal |
| Retorno t-1 | 0.005567 (0.007271) |
| Δ SVI t-1 | -5.52E-05 (0.000327) |
| Ações | (0.000327) |
| Δ SVI t-2 | 0.000258 (0.000410) |
| Ações | (0.000410) |
| Δ SVI t-3 | -0.000356 (0.000432) |
| Ações | (0.000432) |
| Δ SVI t-4 | -0.000541 (0.000411) |
| Ações | (0.000411) |
| Δ SVI t-5 | -0.000413 (0.000329) |
| Ações | (0.000329) |
| Δ volume t-1 | -0.000269*** (9.81E-05) |
| Constante | 0.005744*** (0.000404) |
| Obs. | 18944 |
| Periodos | 256 |
| R2 ajust. | 0 |

Fonte: Dados de pesquisa (2021).

Os resultados sugerem que as atenções semanais e anormais dos investidores no nome das empresas, defasadas do primeiro ao quinto componente autorregressivo, possuem impacto negativo no retorno semanal dos títulos selecionados, com significâncias aos níveis de 5% e 10%. Embora significativa, o seu efeito no retorno semanal é percentualmente baixo, por exemplo, a soma dos efeitos da atenção semanal no nome da empresa defasada foi de 0,0123 p.p, significando que um aumento de 1 p.p na atenção é seguido de uma queda no retorno de 0,0123 p.p. Isso sugere que quanto mais os investidores procuram informação sobre a empresa, motivados por alguma notícia relevante ou divulgação de resultados, maior será o movimento de venda nas semanas subsequentes. Assim como identificado na sessão anterior, esse comportamento é oposto a hipótese “*price pressure hypothesis*”.

As demais regressões não indicaram significância na relação Google Trends e os retornos semanais.

Diferentemente, da subseção anterior que rejeitou a hipótese 4a e aceitou a hipótese 3a, os resultados das equações 6.3 e 6.4 validam a hipótese 4a e rejeitam a hipótese 3a, uma vez que a atenção em *tickers* não refletiu uma variação no retorno semanal dos títulos.

6.3 ATENÇÃO EM TICKERS/EMPRESAS E O VOLUME ANORMAL

Dedica-se essa subseção ao estudo da relação entre a atenção semanal ou anormal dos investidores nos *tickers* de negociação, nome das empresas e nome das empresas + ações e os volumes anormais de negociação dos títulos de ações.

As regressões da variável dependente volume anormal (*Abn_vol*) são especificadas nas equações 6.5 e 6.6.

$$Abn_vol_t = \alpha + \beta_1 Retorno_{t-1} + \sum_{n=1}^5 \beta_2 \Delta SVI_{t-n} + \beta_3 \Delta volume_{t-1} \quad (6.5)$$

$$Abn_vol_t = \alpha + \beta_1 Retorno_{t-1} + \sum_{n=1}^5 \beta_2 ASVI_{t-n} + \beta_3 \Delta volume_{t-1} \quad (6.6)$$

Os coeficientes de regressão e os erros padrões são das estimativas VAR são expressos na Tabela 26, divididos conforme o parâmetro de atenção dos investidores (semanal ou anormal) e a categoria do termo buscado no Google.

TABELA 26 – Dados em painéis - Atenção em empresas/tickers e volumes anormais

A tabela fornece os resultados das regressões com dados em painéis para o volume anormal de um conjunto de ações sobre valores defasados do retorno semanal da ação (retorno), da frequência de buscas no Google (Δ SVI ou ASVI) e do volume semanal de negociação da ação (Δ volume). As regressões foram estimadas para as categorias de pesquisa ticker, nome das empresas e nome das empresas + "ações". Os coeficientes reportados são seguidos por erros padrão entre parênteses. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| Ticker | | | | Empresa | | | |
|----------------------------|----------------------------|---------------------|---------------------------|-----------------------------|-------------------------|---------------------|----------------------------|
| Volume anormal | | Volume anormal | | Volume anormal | | Volume anormal | |
| Retorno t-1 | -0.035415 (0.063780) | Retorno t-1 | -0.035693 (0.063772) | Retorno t-1 | -0.031724 (0.063780) | Retorno t-1 | -0.033221 (0.063789) |
| Δ SVI t-1 Ticker | 0.036265*** (0.005732) | ASVI t-1 Ticker | 0.038558*** (0.004863) | Δ SVI t-1 Empresa | 0.027386 (0.006392) | ASVI t-1 Empresa | 0.025819*** (0.005358) |
| Δ SVI t-2 Ticker | 0.020359*** (0.006979) | ASVI t-2 Ticker | 0.034310*** (0.005661) | Δ SVI t-2 Empresa | 0.011552 (0.007843) | ASVI t-2 Empresa | 0.016588*** (0.006218) |
| Δ SVI t-3 Ticker | -0.004862 (0.007363) | ASVI t-3 Ticker | 0.018335*** (0.005877) | Δ SVI t-3 Empresa | -0.006532 (0.008234) | ASVI t-3 Empresa | 0.003070 (0.006436) |
| Δ SVI t-4 Ticker | -0.028227*** (0.006909) | ASVI t-4 Ticker | -0.002407 (0.005623) | Δ SVI t-4 Empresa | -0.035177 (0.007837) | ASVI t-4 Empresa | -0.022656*** (0.006225) |
| Δ SVI t-5 Ticker | -0.028869*** (0.005650) | ASVI t-5 Ticker | -0.006927 (0.004810) | Δ SVI t-5 Empresa | -0.034670 (0.006403) | ASVI t-5 Empresa | -0.024153*** (0.005374) |
| Δ volume t-1 | 0.007006*** (0.000860) | Δ volume t-1 | 0.006995*** (0.000860) | Δ volume t-1 | 0.007134 (0.000860) | Δ volume t-1 | 0.007129*** (0.000860) |
| Const | -0.004089 (0.003547) | Const | -0.004458 (0.003539) | Const | -0.004126 (0.003539) | Const | -0.004160 (0.003539) |
| Obs. | 18944 | Obs. | 18944 | Obs. | 18944 | Obs. | 18944 |
| Periodos | 256 | Periodos | 256 | Periodos | 256 | Periodos | 256 |
| R2 ajust. | 0.007071 | R2 ajust. | 0.006719 | R2 ajust. | 0.006187 | R2 ajust. | 0.005956 |
| Empresa + "ações" | | | | | | | |
| Volume anormal | | Volume anormal | | Volume anormal | | Volume anormal | |
| Retorno t-1 | -0.025097 (0.063854) | Retorno t-1 | -0.027878 (0.063851) | | | | |
| Δ SVI t-1 Ações | 0.007847*** (0.002869) | ASVI t-1 Ações | 0.010115*** (0.002481) | | | | |
| Δ SVI t-2 Ações | 0.007188** (0.003600) | ASVI t-2 Ações | 0.012471*** (0.003006) | | | | |
| Δ SVI t-3 Ações | 0.000440 (0.003792) | ASVI t-3 Ações | 0.007494** (0.003138) | | | | |
| Δ SVI t-4 Ações | -0.009123** (0.003608) | ASVI t-4 Ações | -0.001813 (0.003016) | | | | |
| Δ SVI t-5 Ações | -0.006264** (0.002886) | ASVI t-5 Ações | -0.002075 (0.002497) | | | | |
| Δ volume t-1 | 0.007147*** (0.000861) | Δ volume t-1 | 0.007136*** (0.000861) | | | | |
| Constante | -0.004230 (0.003545) | Const | -0.004306 (0.003543) | | | | |
| Obs. | 18944 | Obs. | 18944 | | | | |
| Periodos | 256 | Periodos | 256 | | | | |
| R2 ajust. | 0.003753 | R2 ajust. | 0.004037 | | | | |

Fonte: Dados de pesquisa (2021).

Os resultados indicam que as atenções semanais (ΔSVI) em tickers de negociação defasadas em uma e duas semanas possuem impacto positivo nos volumes anormais das ações das semanas subsequentes, com reversão estatisticamente significativa de comportamento na quarta e quinta semana de defasagem. Em outras palavras, quanto mais próximo da semana de negociação, o maior interesse dos investidores nas ações reflete um maior volume de negociação, uma vez que os investidores transformam sua atenção nos tickers em negociação anormal de compra ou venda. Efeito similar foi verificado ao se analisar a atenção anormal no nome das empresas e a atenção semanal no nome da empresa + a palavra ações.

Já a atenção anormal ($ASVI$) em *tickers* de negociação possuem relação positiva com os volumes anormais de negociação em até três defasagens, significando que uma maior atenção anormal em códigos de ações prediz um aumento no volume anormal negociado. Nesse caso, porém, não foi verificada uma reversão estatisticamente significativa no comportamento da atenção anormal defasada em quatro e cinco períodos.

Apenas para a relação entre atenção semanal no nome das empresas e os volumes anormais não foi verificado comportamento estaticamente significativo. Dessa forma, é possível aceitar as hipóteses H3b e H4b no que tange volume de negociação anormal.

6.4 ATENÇÃO EM TICKERS/EMPRESAS E O VOLUME SEMANAL

Complementando a análise da subseção anterior, investigamos a relação da atenção e os volumes semanais de negociação. As regressões da variável dependente (volume semanal) são especificadas nas equações 6.7 e 6.8 abaixo.

$$\Delta \text{volume}_t = \alpha + \beta_1 \text{Retorno}_{t-1} + \sum_{n=1}^5 \beta_2 \Delta SVI_{t-n} + \beta_3 \Delta \text{volume}_{t-1} \quad (6.7)$$

$$\Delta \text{volume}_t = \alpha + \beta_1 \text{Retorno}_{t-1} + \sum_{n=1}^5 \beta_2 ASVI_{t-n} + \beta_3 \Delta \text{volume}_{t-1} \quad (6.8)$$

Os coeficientes de regressão e os erros padrões das estimativas VAR são expressos na Tabela 27, divididos conforme o parâmetro de atenção dos investidores (semanal ou anormal) e a categoria do termo buscado no Google.

TABELA 27 – Dados em painéis - Atenção em empresas/tickers e volumes semanais

A tabela fornece os resultados das regressões com dados em painéis para o volume semanal de um conjunto de ações sobre valores defasados do retorno semanal da ação (retorno), da frequência de buscas no Google (Δ SVI ou ASVI) e do volume semanal de negociação da ação (Δ volume). As regressões foram estimadas para as categorias de pesquisa ticker, nome das empresas e nome das empresas + "ações". Os coeficientes reportados são seguidos por erros padrão entre parênteses. A significância nos níveis de 1%, 5% e 10% é denotada por ***, **, *, respectivamente.

| Ticker | | Empresa | | | | | |
|---------------------|----------------------------|---------------------|----------------------------|---------------------|----------------------------|---------------------|----------------------------|
| Volume semanal | | Volume semanal | | Volume semanal | | Volume semanal | |
| Retorno t-1 | -0.660670 (0.539883) | Retorno t-1 | -0.681399 (0.539769) | Retorno t-1 | -0.723098 (0.539696) | Retorno t-1 | -0.720680 (0.539725) |
| Δ SVI t-1 | -0.106390** (0.048519) | ASVI t-1 | -0.064584 (0.041156) | Δ SVI t-1 | -0.037725 (0.054087) | ASVI t-1 | -0.011963 (0.045335) |
| Ticker | | Ticker | | Empresa | | Empresa | |
| Δ SVI t-2 | -0.109161* (0.059075) | ASVI t-2 | -0.053441 (0.047913) | Δ SVI t-2 | -0.079907 (0.066366) | ASVI t-2 | -0.050615 (0.052615) |
| Ticker | | Ticker | | Empresa | | Empresa | |
| Δ SVI t-3 | -0.093010 (0.062323) | ASVI t-3 | -0.036943 (0.049744) | Δ SVI t-3 | -0.063435 (0.069676) | ASVI t-3 | -0.031393 (0.054459) |
| Ticker | | Ticker | | Empresa | | Empresa | |
| Δ SVI t-4 | -0.069091 (0.058480) | ASVI t-4 | -0.024283 (0.047595) | Δ SVI t-4 | -0.062075 (0.066312) | ASVI t-4 | -0.031471 (0.052672) |
| Ticker | | Ticker | | Empresa | | Empresa | |
| Δ SVI t-5 | -0.078746* (0.047823) | ASVI t-5 | -0.048755 (0.040709) | Δ SVI t-5 | -0.076444 (0.054179) | ASVI t-5 | -0.050314 (0.045469) |
| Ticker | | Ticker | | Empresa | | Empresa | |
| Δ volume t-1 | -0.020693*** (0.007282) | Δ volume t-1 | -0.020761*** (0.007283) | Δ volume t-1 | -0.021074*** (0.007280) | Δ volume t-1 | -0.021083*** (0.007280) |
| Const | 0.188474*** (0.030021) | Const | 0.184570*** (0.029951) | Const | 0.184365*** (0.029947) | Const | 0.183994*** (0.029946) |
| Obs. | 18944 | Obs. | 18944 | Obs. | 18944 | Obs. | 18944 |
| Periodos | 256 | Periodos | 256 | Periodos | 256 | Periodos | 256 |
| R2 ajust. | 0.006835 | R2 ajust. | 0.006655 | R2 ajust. | 0.006614 | R2 ajust. | 0.006577 |
| Empresa + "ações" | | | | | | | |
| Volume semanal | | Volume semanal | | Volume semanal | | Volume semanal | |
| Retorno t-1 | -0.700172 (0.539694) | Retorno t-1 | -0.704293 (0.539750) | | | | |
| Δ SVI t-1 | -0.019115 (0.024250) | ASVI t-1 | -0.007857 (0.020972) | | | | |
| Ações | | Ações | | | | | |
| Δ SVI t-2 | -0.018329 (0.030431) | ASVI t-2 | -0.004103 (0.025408) | | | | |
| Ações | | Ações | | | | | |
| Δ SVI t-3 | -0.019483 (0.032051) | ASVI t-3 | -0.004955 (0.026530) | | | | |
| Ações | | Ações | | | | | |
| Δ SVI t-4 | -0.019778 (0.030491) | ASVI t-4 | -0.007728 (0.025491) | | | | |
| Ações | | Ações | | | | | |
| Δ SVI t-5 | -0.007626 (0.024395) | ASVI t-5 | -0.003480 (0.021109) | | | | |
| Ações | | Ações | | | | | |
| Δ volume t-1 | -0.021052*** (0.007280) | Δ volume t-1 | -0.021074*** (0.007280) | | | | |
| Constante | 0.184532*** (0.029959) | Const | 0.183965*** (0.029948) | | | | |
| Obs. | 18944 | Obs. | 18944 | | | | |
| Periodos | 256 | Periodos | 256 | | | | |
| R2 ajust. | 0.006516 | R2 ajust. | 0.006484 | | | | |

Fonte: Dados de pesquisa (2021).

A variação semanal da frequência de buscas dos *tickers* de negociação defasada em uma, duas e cinco semanas, prediz um impacto negativo no volume de negociação dos títulos, com significância de 5% e 10%. Isto sugere que ao obterem maiores informações sobre as empresas, via *tickers*, os investidores diminuem o fluxo de investimento no mercado, sejam por movimentos compradores ou vendedores, reduzindo o volume transacionado. Ressaltando que o somatório dos coeficientes de atenção defasados foi igual a 0,45, ou seja, um aumento de 1 p.p. na frequência semanal de buscas resulta em uma diminuição no volume semanal negociado em 0,45 p.p.

As demais regressões não indicaram significância na relação Google Trends e os volumes semanais.

Concluindo, os resultados dessa subseção de volumes semanais validam a hipótese H3b, relacionada aos *tickers* de negociação e o volume semanal, mas rejeitam a hipótese H4b, relacionada ao nome das empresas.

7 ESTRATÉGIAS DE INVESTIMENTOS

Nesta seção exploramos os resultados dos modelos econométricos, avaliando seus desempenhos em uma estratégia de alocação de capitais, desenvolvida com base em inteligência artificial de redes neurais.

7.1 ESTRUTURAÇÃO REDE NEURAL ARTIFICIAL (RNA)

O objetivo de uma rede neural artificial (RNA), conforme Rich e Knight (1993), é simular o funcionamento do cérebro humano via utilização de neurônios artificiais em métodos computacionais programados. Complementando, Haykin (2001), sugere que uma rede neural possui a capacidade de armazenar conhecimentos experimentais e transformá-los em informações disponíveis para o uso e, conseqüentemente, melhorar a assertividade nas tomadas de decisões.

O propósito da rede neural do presente trabalho é prever possíveis valorizações semanais do índice Ibovespa, com base em dados extraídos do Google Trends. Dessa forma, espera-se auxiliar possíveis investidores em seus direcionamentos de alocação de capital.

Uma rede de neurônios pode ser organizada em diversas topologias e arquiteturas. Conforme Neto (2014), o problema a ser tratado é que vai direcionar qual é o melhor tipo de rede a ser utilizada, podendo ser classificada conforme número de camadas ocultas (*single layer ou multilayer*), número de neurônios por camada oculta e quanto ao sentido do fluxo de dados (*feedforward ou recorrente*). Outra característica importante para a formação das redes neurais é a função de ativação dos neurônios, que limita a faixa de valores de saída do neurônio artificial, podendo ser classificada entre linear, degrau, tansig, logsig e entre outros.

A análise proposta utilizou a arquitetura da rede neural alimentada diretamente em uma direção (*feedforward*) com duas camadas ocultas, cada uma com número de neurônios variando entre 1 e 20 e funções de ativação logsig e tansig. Modelo similar foi proposto por Correia (2015) e Almeida (2013), ao analisarem a valorização ou desvalorização de IPOs no primeiro dia de negociação.

O grafo na Figura 6 ilustra a estrutura da arquitetura selecionada, sendo que, nessa organização estrutural, segundo Neto (2014, pág. 50), “(...) além de haver um processamento paralelo entre os neurônios dentro de uma mesma camada, há um processamento em série com um que percorre uma ou mais camadas até ser alcançada a camada de saída”

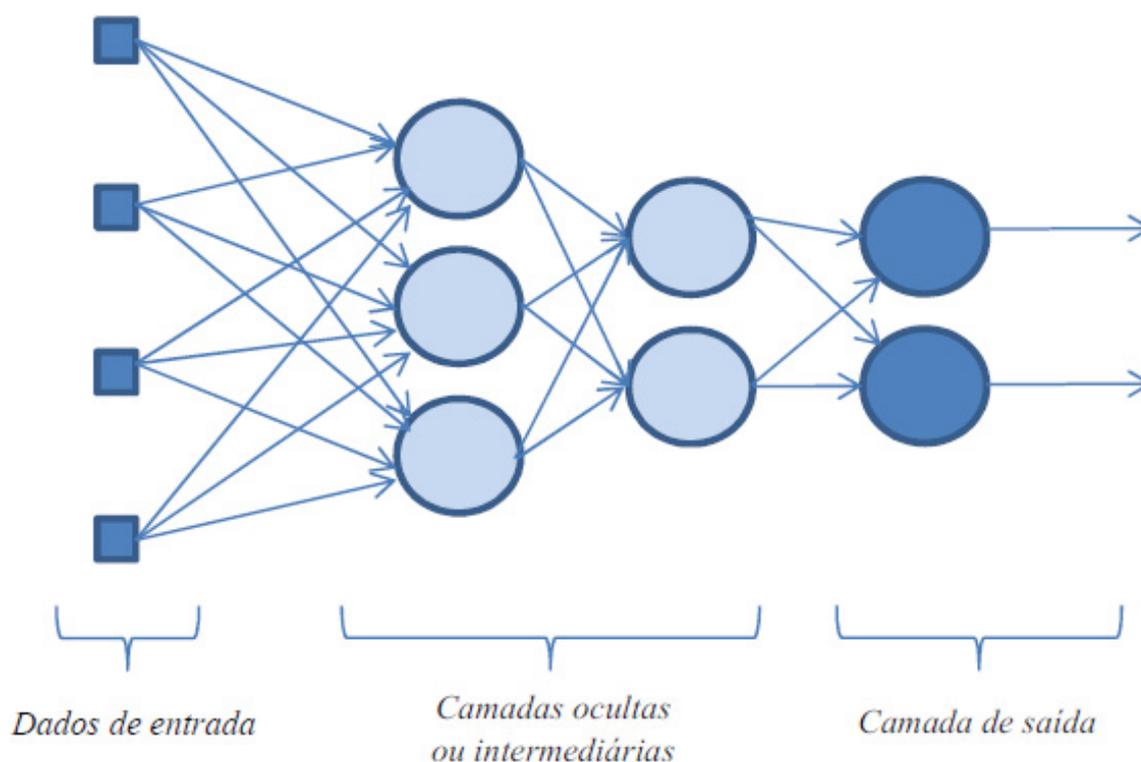


FIGURA 6 – Estrutura de uma Rede Neural com várias camadas.
 FONTE: NETO (2014)

7.2 PARÂMETROS DA REDE NEURAL ARTIFICIAL

Como dados de entrada do modelo, utilizamos os índices de sentimentos FEARS positivo e negativo, posto que tiveram relação estatisticamente significativa no retorno semanal do índice Ibovespa. Além disso, também utilizamos as variações das frequências semanais de buscas da palavra “Ibovespa”, uma vez que a variação anormal dessa variável teve relação significativa com o retorno semanal do Ibovespa.

O único parâmetro de saída da rede foi a variação semanal do retorno do índice Ibovespa, onde caso a variação tenha sido positiva, o valor 1 é atribuído a

variável de saída, caso essa variação tenha sido negativa, o valor -1 é atribuído a variável de saída. Os parâmetros analisados abrangeram o período entre 28/06/2015 e 10/05/2019.

Martineli (1999, pág 44) afirmou que para obtenção de um resultado satisfatório de uma Rede Neural, “(...) devem ser feitas várias divisões do mesmo conjunto de dados em conjunto de treinamento e de testes. Este método, chamado de *cross validation*, fornece uma estimativa mais realista do erro a ser cometido pelo método de classificação utilizado”. Na mesma linha, Barone (2003) separou a operação de desenvolvimento das RNAs em duas fases:

1. Treinamento: quando os valores das sinapses estão sendo ajustados.
2. Atuação ou teste: quando os valores das sinapses não mudam.

Ainda conforme Martineli (1999), o conjunto de treinamento utiliza mais exemplos de uma classe em relação ao conjunto de teste. Portanto, é necessária uma distribuição de classes balanceada, a fim de favorecer uma análise estatística mais confiável. O presente trabalho optou por dividir a amostra na proporção de 70% para treinamento e 30% para teste, seguindo a porcentagem aplicada no trabalho de Roque (2009). O Quadro 7 apresenta um resumo das principais informações do modelo.

QUADRO 7 – Características do conjunto de dados para análise.

| | |
|---|--|
| Parâmetros de entrada | Variação da frequência semanal de buscas da palavra “Ibovespa” defasados em 1 a 5 períodos. Índices de sentimento FEARS positivo e negativo, defasados em 1 a 5 períodos. |
| Parâmetros de saída | Variação semanal do retorno do índice Ibovespa |
| Período analisado | 28/06/2015 a 10/05/2019 |
| Número de semanas | 202 |
| Nº de semanas selecionadas para treinamento (70%) | 141 |
| Nº de semanas selecionadas para teste (30%) | 61 |
| Número de colunas de entrada na matriz | 15 |

FONTE: O autor (2021)

7.3 DESENVOLVIMENTO DA REDE NEURAL ARTIFICIAL NO MATLAB

As Redes Neurais Artificiais foram desenvolvidas e executadas no software MATLAB®, conforme script apresentado na Figura 7, as explicações das principais linhas do código são realizadas na sequência. A partir dos resultados obtidos, a análise dos dados e a elaboração das estratégias de investimento foram realizadas em planilha eletrônica.

```

1 - A=load('Teste_2.txt');
2 - size(A)
3 - pause
4 - P1=A(1:179,1:10)';
5 - P2=A(180:256,1:10)';
6 - T1=A(1:179,11)';
7 - T2=A(180:256,11)';
8 - pause
9 - for i=1:20
10 -     for j=1:20
11 -         net=feedforwardnet(P1,T1,[i j],{'tansig' 'logsig'},'traingd')
12 -         net=train(net,P1,T1);
13 -         X((i-1)*20+j,:)=sim(net,P1);
14 -         Y((i-1)*20+j,:)=sim(net,P2);
15 -     end
16 - end
17 - for i=1:20
18 -     for j=1:20
19 -         net=feedforwardnet(P1,T1,[i j],{'tansig' 'tansig'},'traingd')
20 -         net=train(net,P1,T1);
21 -         X(400+(i-1)*20+j,:)=sim(net,P1);
22 -         Y(400+(i-1)*20+j,:)=sim(net,P2);
23 -     end
24 - end
25 - for i=1:20
26 -     for j=1:20
27 -         net=feedforwardnet(P1,T1,[i j],{'logsig' 'tansig'},'traingd')
28 -         net=train(net,P1,T1);
29 -         X(800+(i-1)*20+j,:)=sim(net,P1);
30 -         Y(800+(i-1)*20+j,:)=sim(net,P2);
31 -     end
32 - end
33 - for i=1:20
34 -     for j=1:20
35 -         net=feedforwardnet(P1,T1,[i j],{'logsig' 'logsig'},'traingd')
36 -         net=train(net,P1,T1);
37 -         X(1200+(i-1)*20+j,:)=sim(net,P1);
38 -         Y(1200+(i-1)*20+j,:)=sim(net,P2);
39 -     end
40 - end
41 - save 'resultadoTestellY.txt' Y -ascii
42 - save 'resultadoTestellX.txt' X -ascii

```

FIGURA 7 – Script da rede neural no software MATLAB®

FONTE: O AUTOR (2021) com base no trabalho de ALMEIDA (2013) e CORREIA (2015)

A linha 1 representa o comando para inserir a matriz de dados e defini-la como uma variável "A", essa matriz de dados é carregada através do bloco de notas e de um arquivo .txt.

As linhas de 4 a 7 definem qual será o tamanho da matriz de treinamento e da matriz de teste, a nossa rede utilizou uma proporção de 70%-30%. Adicionalmente, a matriz P1 contém as entradas dos dados utilizados para treinamento, a matriz P2, contém as entradas dos dados utilizados para teste, a matriz T1, contém as respostas esperadas dos dados de treinamento e a matriz T2 contém as repostas esperadas dos dados de teste. Como existe uma resposta esperada, a aprendizagem desse teste é supervisionada.

Em seguida, na linha 11, a rede de camadas múltiplas é iniciada pela função *feedforwardnet*, tal função foi utilizada com cinco parâmetros de entrada, são eles, conforme Almeida (2013), [1] O primeiro argumento é a matriz P1, que contém os parâmetros de entrada dos dados de treinamento; [2] O segundo argumento é a matriz T1, que contém as respostas dos dados de treinamento; [3] O terceiro argumento é um vetor que contém o número de neurônios de cada camada oculta; [4] O quarto argumento são as funções de ativação para cada camada oculta (*tansig* ou *logsig*); [5] O último argumento contém o algoritmo de treinamento utilizado (*traingd*).

Na linha 12, o treinamento da rede é realizado através da utilização da função NET. Tal função define qual rede ira ser treinada, a matriz P1 com os parâmetros de entrada e a matriz T1.

Após o treinamento, de acordo com Almeida (2013), a função sim da linha 13, realiza uma simulação no próprio software com os dados de teste contidos na matriz de treinamento P1, a variável X vai sendo carregada com os resultados obtidos nessas simulações. O mesmo ocorre para a função da linha 13, quando a variável Y vai sendo carregada pelas simulações da matriz de treinamento P2.

Os códigos entre as linhas 9 e 16 são repetidos mais 3 vezes, alterando conforme a função de ativação de cada uma das camadas ocultas.

Os resultados dos conjuntos de treinamento e teste das Redes são extraídos através dos códigos das linhas 41 e 42. Após executar o programa, espera-se obter a valorização ou desvalorização do título no período proposto, para cada uma das 1600 redes treinadas.

7.4 ANÁLISE DOS RESULTADOS

Nesta etapa da pesquisa, o objetivo é avaliar os resultados obtidos nas etapas de treinamento e teste para cada uma das redes arquitetadas. É importante ressaltar que a análise foi realizada buscando avaliar se as configurações de rede foram capazes de prever se o índice Ibovespa iria se valorizar ou desvalorizar ao longo de cada uma das semanas, ou seja, não foram analisados os percentuais de rentabilidade sugeridos pelas redes neurais.

Em razão dos neurônios nas camadas escondidas variarem de 1 até 20 e as funções de ativação variarem entre *logsig* e *tansig*, foram executadas um total de 1.600 configurações de rede. Primeiramente, calculamos os percentuais de acerto no treinamento para cada uma das 1.600 configurações de rede, as 5 configurações que tiveram melhor desempenho foram selecionadas e o percentual de acerto de cada uma dessas configurações foi calculado com base nos dados de teste.

Isto posto, das 1.600 configurações de rede desenvolvidas, 1.401 obtiveram mais acertos do que erros em relação ao resultado esperado durante a etapa de treinamento, representando cerca de 87,5% de acerto, sendo que as 5 redes com maior percentual de acerto podem ser observadas na Tabela 28, além disso é possível visualizar os resultados dessas configurações durante a etapa de testes.

TABELA 28 – Desempenho das 05 melhores configurações de rede

| Configuração De Rede | TREINAMENTO | | | TESTE | | |
|----------------------|------------------|---------------|--------|------------------|---------------|--------|
| | Total de semanas | Nº de acertos | % | Total de semanas | Nº de acertos | % |
| 717 | 141 | 109 | 77,30% | 61 | 36 | 59,02% |
| 480 | 141 | 109 | 77,30% | 61 | 34 | 55,74% |
| 656 | 141 | 108 | 76,60% | 61 | 38 | 62,30% |
| 255 | 141 | 107 | 75,89% | 61 | 34 | 55,74% |
| 813 | 141 | 107 | 75,89% | 61 | 26 | 42,62% |

FONTE: O autor (2021)

Conforme apresentado na Tabela 28, verificamos que a rede 717 e 480 foram as que obtiveram o melhor desempenho no treinamento (77,30%). Apesar do percentual de acerto ser reduzido na etapa de testes em relação a etapa de treinamento, ainda sim, ambas as redes obtiveram taxa de acerto superior a 50%. Com base nos resultados da rede 717, o Quadro 8 apresenta a comparação entre as

decisões de compra ou não compra sugeridas pelo modelo e o retorno real do índice Ibovespa, indicando os 59,02% de acerto.

Dentre as 5 redes com melhor desempenho no treinamento, a rede 656 apresentou a maior taxa de acerto na fase de testes, resultando em um percentual de acerto de 62,30%, ou seja, dentre as 61 semanas analisadas, o modelo acertou a decisão de compra ou não compra em 36 semanas.

QUADRO 8 – Resultado do teste da melhor configuração de rede.

| Semana | Decisão do modelo Rede Neural | Retorno semanal Ibovespa | Resultado da Rede |
|------------|-------------------------------|--------------------------|-------------------|
| 11/03/2018 | Não Compra | Desvalorizou | Acertou |
| 18/03/2018 | Não Compra | Desvalorizou | Acertou |
| 25/03/2018 | Compra | Valorizou | Acertou |
| 01/04/2018 | Não Compra | Desvalorizou | Acertou |
| 08/04/2018 | Compra | Desvalorizou | Errou |
| 15/04/2018 | Não Compra | Valorizou | Errou |
| 22/04/2018 | Compra | Valorizou | Acertou |
| 29/04/2018 | Não Compra | Desvalorizou | Acertou |
| 06/05/2018 | Compra | Valorizou | Acertou |
| 13/05/2018 | Compra | Desvalorizou | Errou |
| 20/05/2018 | Não Compra | Desvalorizou | Acertou |
| 27/05/2018 | Compra | Desvalorizou | Errou |
| 03/06/2018 | Não Compra | Desvalorizou | Acertou |
| 10/06/2018 | Não Compra | Desvalorizou | Acertou |
| 17/06/2018 | Compra | Desvalorizou | Errou |
| 24/06/2018 | Compra | Valorizou | Acertou |
| 01/07/2018 | Compra | Valorizou | Acertou |
| 08/07/2018 | Compra | Valorizou | Acertou |
| 15/07/2018 | Não Compra | Valorizou | Errou |
| 22/07/2018 | Não Compra | Valorizou | Errou |
| 29/07/2018 | Compra | Valorizou | Acertou |
| 05/08/2018 | Compra | Desvalorizou | Errou |
| 12/08/2018 | Não Compra | Desvalorizou | Acertou |
| 19/08/2018 | Compra | Valorizou | Acertou |
| 26/08/2018 | Compra | Valorizou | Acertou |
| 02/09/2018 | Compra | Desvalorizou | Errou |
| 09/09/2018 | Compra | Desvalorizou | Errou |
| 16/09/2018 | Compra | Valorizou | Acertou |
| 23/09/2018 | Compra | Desvalorizou | Errou |
| 30/09/2018 | Compra | Valorizou | Acertou |
| 07/10/2018 | Compra | Valorizou | Acertou |
| 14/10/2018 | Não Compra | Valorizou | Errou |
| 21/10/2018 | Não Compra | Valorizou | Errou |
| 28/10/2018 | Não Compra | Valorizou | Errou |
| 04/11/2018 | Compra | Desvalorizou | Errou |
| 11/11/2018 | Compra | Valorizou | Acertou |
| 18/11/2018 | Compra | Desvalorizou | Errou |
| 25/11/2018 | Não Compra | Valorizou | Errou |
| 02/12/2018 | Não Compra | Desvalorizou | Acertou |
| 09/12/2018 | Não Compra | Desvalorizou | Acertou |
| 16/12/2018 | Compra | Desvalorizou | Errou |
| 23/12/2018 | Compra | Valorizou | Acertou |
| 30/12/2018 | Compra | Valorizou | Acertou |
| 06/01/2019 | Não Compra | Valorizou | Errou |
| 13/01/2019 | Compra | Valorizou | Acertou |
| 20/01/2019 | Compra | Valorizou | Acertou |
| 27/01/2019 | Compra | Valorizou | Acertou |
| 03/02/2019 | Não Compra | Desvalorizou | Acertou |

| Semana | Decisão do modelo Rede Neural | Retorno semanal Ibovespa | Resultado da Rede |
|------------|-------------------------------|--------------------------|-------------------|
| 10/02/2019 | Não Compra | Valorizou | Errou |
| 17/02/2019 | Compra | Valorizou | Acertou |
| 24/02/2019 | Compra | Desvalorizou | Errou |
| 03/03/2019 | Compra | Valorizou | Acertou |
| 10/03/2019 | Não Compra | Valorizou | Errou |
| 17/03/2019 | Não Compra | Desvalorizou | Acertou |
| 24/03/2019 | Compra | Valorizou | Acertou |
| 31/03/2019 | Não Compra | Valorizou | Errou |
| 07/04/2019 | Compra | Desvalorizou | Errou |
| 14/04/2019 | Não Compra | Valorizou | Errou |
| 21/04/2019 | Compra | Valorizou | Acertou |
| 28/04/2019 | Não Compra | Desvalorizou | Acertou |
| 05/05/2019 | Não Compra | Desvalorizou | Acertou |

FONTE: O autor (2021)

Por fim, o Quadro 9 apresenta qual é a função de ativação por camada oculta e qual a quantidade de neurônios ativados por camada.

QUADRO 9 - Características das 05 melhores configurações de redes

| Configuração De Rede | 1ª Camada Oculta | | 2ª Camada Oculta | |
|----------------------|--------------------|-----------------|--------------------|-----------------|
| | Função De Ativação | Nº De Neurônios | Função De Ativação | Nº De Neurônios |
| 717 | tansig | 16 | tansig | 17 |
| 480 | tansig | 4 | tansig | 20 |
| 656 | tansig | 13 | tansig | 16 |
| 255 | tansig | 13 | logsig | 15 |
| 813 | logsig | 1 | tansig | 13 |

FONTE: O autor (2021)

7.5 SIMULAÇÃO DE INVESTIMENTO

A partir de duas simulações de investimentos, comparamos as rentabilidades das estratégias baseadas em um modelo desenvolvido em redes neurais e informações extraídas do Google Trends com a rentabilidade do índice Ibovespa no mesmo período. É importante ressaltar que os custos de operação (corretagem, impostos e custódia) não foram considerados nas análises.

7.5.1 Simulação de investimento 1

Para a primeira simulação de investimentos, consideramos que o investidor teria à sua disposição R\$ 10.000,00 para aplicar nos inícios das semanas em que a rede neural 717 (com o melhor desempenho de treinamento) sugerisse a compra do índice Ibovespa. Ao término de cada semana de compra, o investidor realizaria a operação de venda do montante acumulado. Os resultados foram comparados com a estratégia de comprar o índice no início de todas as semanas, no valor de R\$ 10.000,00, e vendê-lo no fechamento semanal. O período de abrangência da análise foi o período semanal entre 11/03/2018 e 05/05/2019, ou seja, as semanas selecionadas para treinamento da rede neural.

Com base no Quadro 10, o modelo proposto pela rede neural obteve um retorno acumulado de R\$ 1.655,00, enquanto a estratégia de compra e venda em todas as semanas obteve retorno acumulado de R\$ 1.092,00, resultando em uma diferença de R\$ 563,00.

Portanto, para esse exemplo, ao se comparar os dois resultados obtidos, verificamos que o investimento baseado em redes neurais e frequências de busca do Google apresentou um melhor resultado em comparação com o investimento de se comprar e vender todas as semanas, dessa forma, validando positivamente o modelo sugerido.

QUADRO 10 – Resultado simulação de investimentos 1

| Semana | Decisão do modelo Rede Neural | Retorno semanal Ibovespa | Retorno decisão Rede Neural (R\$) | Retorno Ibovespa (R\$) | Diferença (R\$) |
|--------------|-------------------------------|--------------------------|-----------------------------------|------------------------|-----------------|
| 11/03/2018 | Não Compra | (1,72%) | 0,00 | (171,93) | 171,93 |
| 18/03/2018 | Não Compra | (0,60%) | 0,00 | (59,96) | 59,96 |
| 25/03/2018 | Compra | 1,17% | 117,21 | 117,21 | 0,00 |
| 01/04/2018 | Não Compra | (0,64%) | 0,00 | (63,96) | 63,96 |
| 08/04/2018 | Compra | (0,57%) | (57,30) | (57,30) | 0,00 |
| 15/04/2018 | Não Compra | 1,44% | 0,00 | 144,19 | (144,19) |
| 22/04/2018 | Compra | 1,05% | 104,62 | 104,62 | 0,00 |
| 29/04/2018 | Não Compra | (3,85%) | 0,00 | (384,87) | 384,87 |
| 06/05/2018 | Compra | 2,53% | 252,89 | 252,89 | 0,00 |
| 13/05/2018 | Compra | (2,51%) | (250,88) | (250,88) | 0,00 |
| 20/05/2018 | Não Compra | (5,04%) | 0,00 | (503,60) | 503,60 |
| 27/05/2018 | Compra | (2,10%) | (210,14) | (210,14) | 0,00 |
| 03/06/2018 | Não Compra | (5,56%) | 0,00 | (556,45) | 556,45 |
| 10/06/2018 | Não Compra | (2,99%) | 0,00 | (299,42) | 299,42 |
| 17/06/2018 | Compra | (0,17%) | (16,54) | (16,54) | 0,00 |
| 24/06/2018 | Compra | 3,00% | 300,39 | 300,39 | 0,00 |
| 01/07/2018 | Compra | 3,09% | 308,81 | 308,81 | 0,00 |
| 08/07/2018 | Compra | 2,11% | 211,17 | 211,17 | 0,00 |
| 15/07/2018 | Não Compra | 2,58% | 0,00 | 258,11 | (258,11) |
| 22/07/2018 | Não Compra | 1,65% | 0,00 | 164,82 | (164,82) |
| 29/07/2018 | Compra | 1,96% | 196,45 | 196,45 | 0,00 |
| 05/08/2018 | Compra | (6,04%) | (604,29) | (604,29) | 0,00 |
| 12/08/2018 | Não Compra | (0,63%) | 0,00 | (63,39) | 63,39 |
| 19/08/2018 | Compra | 0,31% | 30,65 | 30,65 | 0,00 |
| 26/08/2018 | Compra | 0,55% | 54,55 | 54,55 | 0,00 |
| 02/09/2018 | Compra | (0,34%) | (34,17) | (34,17) | 0,00 |
| 09/09/2018 | Compra | (1,29%) | (129,16) | (129,16) | 0,00 |
| 16/09/2018 | Compra | 5,32% | 532,29 | 532,29 | 0,00 |
| 23/09/2018 | Compra | (0,13%) | (12,84) | (12,84) | 0,00 |
| 30/09/2018 | Compra | 3,76% | 375,59 | 375,59 | 0,00 |
| 07/10/2018 | Compra | 0,73% | 72,76 | 72,76 | 0,00 |
| 14/10/2018 | Não Compra | 1,57% | 0,00 | 156,66 | (156,66) |
| 21/10/2018 | Não Compra | 1,78% | 0,00 | 178,10 | (178,10) |
| 28/10/2018 | Não Compra | 3,15% | 0,00 | 314,86 | (314,86) |
| 04/11/2018 | Compra | (3,14%) | (314,19) | (314,19) | 0,00 |
| 11/11/2018 | Compra | 3,36% | 335,59 | 335,59 | 0,00 |
| 18/11/2018 | Compra | (2,58%) | (258,15) | (258,15) | 0,00 |
| 25/11/2018 | Não Compra | 3,80% | 0,00 | 379,68 | (379,68) |
| 02/12/2018 | Não Compra | (1,55%) | 0,00 | (155,19) | 155,19 |
| 09/12/2018 | Não Compra | (0,75%) | 0,00 | (75,47) | 75,47 |
| 16/12/2018 | Compra | (2,00%) | (200,46) | (200,46) | 0,00 |
| 23/12/2018 | Compra | 2,56% | 255,55 | 255,55 | 0,00 |
| 30/12/2018 | Compra | 4,50% | 449,90 | 449,90 | 0,00 |
| 06/01/2019 | Não Compra | 1,98% | 0,00 | 197,84 | (197,84) |
| 13/01/2019 | Compra | 2,60% | 260,42 | 260,42 | 0,00 |
| 20/01/2019 | Compra | 1,64% | 164,42 | 164,42 | 0,00 |
| 27/01/2019 | Compra | 0,19% | 18,84 | 18,84 | 0,00 |
| 03/02/2019 | Não Compra | (2,57%) | 0,00 | (257,30) | 257,30 |
| 10/02/2019 | Não Compra | 2,29% | 0,00 | 228,96 | (228,96) |
| 17/02/2019 | Compra | 0,37% | 36,91 | 36,91 | 0,00 |
| 24/02/2019 | Compra | (3,35%) | (335,29) | (335,29) | 0,00 |
| 03/03/2019 | Compra | 0,80% | 80,44 | 80,44 | 0,00 |
| 10/03/2019 | Não Compra | 3,96% | 0,00 | 395,53 | (395,53) |
| 17/03/2019 | Não Compra | (5,45%) | 0,00 | (544,90) | 544,90 |
| 24/03/2019 | Compra | 1,79% | 179,23 | 179,23 | 0,00 |
| 31/03/2019 | Não Compra | 1,77% | 0,00 | 177,44 | (177,44) |
| 07/04/2019 | Compra | (4,36%) | (435,91) | (435,91) | 0,00 |
| 14/04/2019 | Não Compra | 1,83% | 0,00 | 183,36 | (183,36) |
| 21/04/2019 | Compra | 1,75% | 175,31 | 175,31 | 0,00 |
| 28/04/2019 | Não Compra | (0,24%) | 0,00 | (23,69) | 23,69 |
| 05/05/2019 | Não Compra | (1,82%) | 0,00 | (182,28) | 182,28 |
| TOTAL | | | 1.654,7 | 1.091,8 | 562,8 |

FONTE: O autor (2021)

7.5.2 Simulação do investimento 2

Na segunda simulação realizada consideramos que o investidor teria à sua disposição R\$ 10.000,00 para aplicar no início da série temporal. No início de cada semana, o montante investido seria a soma do montante inicial mais a rentabilidade acumulada nas semanas anteriores. Ao término de cada semana de compra, o investidor realizaria a operação de venda do montante acumulado.

Com base nos resultados apresentados no Quadro 10, o modelo proposto pela rede neural acumulou um total de R\$ 11.662,00, representando um aumento de capital de 17% ou R\$ 1.662,00, enquanto a estratégia de *buy and hold* acumulou um total de R\$ 10.913,00, representando um aumento de capital de 9% ou R\$ 913,00. Embora ambos os retornos sejam positivos, o investimento baseado em Redes Neurais apresentou um resultado 82% superior a estratégia *buy and hold*.

Assim como na primeira simulação realizada, os resultados validaram positivamente o modelo de rede neural sugerido.

QUADRO 11 – Resultado simulação de investimentos 2.

| Semana | Decisão do modelo Rede Neural | Retorno semanal Ibovespa | Retorno decisão Rede Neural (R\$) | Saldo Rede Neural (R\$) | Retorno Ibovespa (R\$) | Saldo IBOV (R\$) | Diferença (R\$) |
|--------------|-------------------------------|--------------------------|-----------------------------------|-------------------------|------------------------|------------------|-----------------|
| 11/03/2018 | Não Compra | (1,72%) | 0,00 | 10.000 | (171,93) | 9.828 | 171,93 |
| 18/03/2018 | Não Compra | (0,60%) | 0,00 | 10.000 | (58,93) | 9.769 | 230,86 |
| 25/03/2018 | Compra | 1,17% | 117,21 | 10.117 | 114,51 | 9.884 | 233,57 |
| 01/04/2018 | Não Compra | (0,64%) | 0,00 | 10.117 | (63,22) | 9.820 | 296,79 |
| 08/04/2018 | Compra | (0,57%) | (57,97) | 10.059 | (56,27) | 9.764 | 295,09 |
| 15/04/2018 | Não Compra | 1,44% | 0,00 | 10.059 | 140,79 | 9.905 | 154,30 |
| 22/04/2018 | Compra | 1,05% | 105,24 | 10.164 | 103,62 | 10.009 | 155,91 |
| 29/04/2018 | Não Compra | (3,85%) | 0,00 | 10.164 | (385,20) | 9.623 | 541,11 |
| 06/05/2018 | Compra | 2,53% | 257,05 | 10.422 | 243,37 | 9.867 | 554,79 |
| 13/05/2018 | Compra | (2,51%) | (261,46) | 10.160 | (247,54) | 9.619 | 540,88 |
| 20/05/2018 | Não Compra | (5,04%) | 0,00 | 10.160 | (484,42) | 9.135 | 1.025,30 |
| 27/05/2018 | Compra | (2,10%) | (213,51) | 9.947 | (191,96) | 8.943 | 1.003,75 |
| 03/06/2018 | Não Compra | (5,56%) | 0,00 | 9.947 | (497,62) | 8.445 | 1.501,37 |
| 10/06/2018 | Não Compra | (2,99%) | 0,00 | 9.947 | (252,86) | 8.192 | 1.754,24 |
| 17/06/2018 | Compra | (0,17%) | (16,45) | 9.930 | (13,55) | 8.179 | 1.751,33 |
| 24/06/2018 | Compra | 3,00% | 298,29 | 10.228 | 245,68 | 8.424 | 1.803,94 |
| 01/07/2018 | Compra | 3,09% | 315,86 | 10.544 | 260,16 | 8.685 | 1.859,65 |
| 08/07/2018 | Compra | 2,11% | 222,67 | 10.767 | 183,39 | 8.868 | 1.898,92 |
| 15/07/2018 | Não Compra | 2,58% | 0,00 | 10.767 | 228,90 | 9.097 | 1.670,03 |
| 22/07/2018 | Não Compra | 1,65% | 0,00 | 10.767 | 149,93 | 9.247 | 1.520,09 |
| 29/07/2018 | Compra | 1,96% | 211,52 | 10.978 | 181,66 | 9.429 | 1.549,95 |
| 05/08/2018 | Compra | (6,04%) | (663,41) | 10.315 | (569,75) | 8.859 | 1.456,29 |
| 12/08/2018 | Não Compra | (0,63%) | 0,00 | 10.315 | (56,15) | 8.803 | 1.512,45 |
| 19/08/2018 | Compra | 0,31% | 31,61 | 10.347 | 26,98 | 8.830 | 1.517,08 |
| 26/08/2018 | Compra | 0,55% | 56,44 | 10.403 | 48,16 | 8.878 | 1.525,36 |
| 02/09/2018 | Compra | (0,34%) | (35,55) | 10.368 | (30,33) | 8.847 | 1.520,14 |
| 09/09/2018 | Compra | (1,29%) | (133,91) | 10.234 | (114,27) | 8.733 | 1.500,51 |
| 16/09/2018 | Compra | 5,32% | 544,73 | 10.778 | 464,86 | 9.198 | 1.580,38 |
| 23/09/2018 | Compra | (0,13%) | (13,84) | 10.765 | (11,81) | 9.186 | 1.578,35 |
| 30/09/2018 | Compra | 3,76% | 404,30 | 11.169 | 345,02 | 9.531 | 1.637,63 |
| 07/10/2018 | Compra | 0,73% | 81,27 | 11.250 | 69,35 | 9.601 | 1.649,55 |
| 14/10/2018 | Não Compra | 1,57% | 0,00 | 11.250 | 150,40 | 9.751 | 1.499,15 |
| 21/10/2018 | Não Compra | 1,78% | 0,00 | 11.250 | 173,67 | 9.925 | 1.325,48 |
| 28/10/2018 | Não Compra | 3,15% | 0,00 | 11.250 | 312,49 | 10.237 | 1.012,99 |
| 04/11/2018 | Compra | (3,14%) | (353,46) | 10.897 | (321,64) | 9.915 | 981,16 |
| 11/11/2018 | Compra | 3,36% | 365,68 | 11.262 | 332,75 | 10.248 | 1.014,09 |
| 18/11/2018 | Compra | (2,58%) | (290,73) | 10.972 | (264,56) | 9.984 | 987,91 |
| 25/11/2018 | Não Compra | 3,80% | 0,00 | 10.972 | 379,06 | 10.363 | 608,85 |
| 02/12/2018 | Não Compra | (1,55%) | 0,00 | 10.972 | (160,82) | 10.202 | 769,67 |
| 09/12/2018 | Não Compra | (0,75%) | 0,00 | 10.972 | (76,99) | 10.125 | 846,66 |
| 16/12/2018 | Compra | (2,00%) | (219,93) | 10.752 | (202,96) | 9.922 | 829,69 |
| 23/12/2018 | Compra | 2,56% | 274,76 | 11.026 | 253,56 | 10.176 | 850,89 |
| 30/12/2018 | Compra | 4,50% | 496,07 | 11.522 | 457,79 | 10.633 | 889,17 |
| 06/01/2019 | Não Compra | 1,98% | 0,00 | 11.522 | 210,37 | 10.844 | 678,80 |
| 13/01/2019 | Compra | 2,60% | 300,06 | 11.823 | 282,39 | 11.126 | 696,48 |
| 20/01/2019 | Compra | 1,64% | 194,38 | 12.017 | 182,93 | 11.309 | 707,93 |
| 27/01/2019 | Compra | 0,19% | 22,64 | 12.040 | 21,30 | 11.330 | 709,26 |
| 03/02/2019 | Não Compra | (2,57%) | 0,00 | 12.040 | (291,53) | 11.039 | 1.000,80 |
| 10/02/2019 | Não Compra | 2,29% | 0,00 | 12.040 | 252,75 | 11.292 | 748,05 |
| 17/02/2019 | Compra | 0,37% | 44,44 | 12.084 | 41,68 | 11.333 | 750,81 |
| 24/02/2019 | Compra | (3,35%) | (405,16) | 11.679 | (379,99) | 10.953 | 725,64 |
| 03/03/2019 | Compra | 0,80% | 93,95 | 11.773 | 88,11 | 11.041 | 731,48 |
| 10/03/2019 | Não Compra | 3,96% | 0,00 | 11.773 | 436,72 | 11.478 | 294,75 |
| 17/03/2019 | Não Compra | (5,45%) | 0,00 | 11.773 | (625,44) | 10.853 | 920,20 |
| 24/03/2019 | Compra | 1,79% | 211,00 | 11.984 | 194,51 | 11.047 | 936,69 |
| 31/03/2019 | Não Compra | 1,77% | 0,00 | 11.984 | 196,01 | 11.243 | 740,67 |
| 07/04/2019 | Compra | (4,36%) | (522,38) | 11.461 | (490,10) | 10.753 | 708,39 |
| 14/04/2019 | Não Compra | 1,83% | 0,00 | 11.461 | 197,17 | 10.950 | 511,21 |
| 21/04/2019 | Compra | 1,75% | 200,92 | 11.662 | 191,96 | 11.142 | 520,18 |
| 28/04/2019 | Não Compra | (0,24%) | 0,00 | 11.662 | (26,40) | 11.116 | 546,57 |
| 05/05/2019 | Não Compra | (1,82%) | 0,00 | 11.662 | (202,61) | 10.913 | 749,19 |
| TOTAL | | | 1.662,3 | 11.662 | 913,2 | 10.913 | 749,2 |

FONTE: O autor (2021)

8 CONCLUSÃO

O trabalho teve como objetivo avaliar a relação entre atenção e sentimento dos investidores, calculados com base na frequência de buscas semanais de palavras-chave no Google, e indicadores de desempenho do mercado brasileiro de ações, como volume de negociações, retorno semanal e volatilidade do mercado. O entendimento dessa relação tinha como meta gerar possíveis insights que auxiliem investidores em suas tomadas de decisões em busca de uma maior rentabilidade de seus investimentos e um menor risco associado, desafiando os pressupostos centrais da Hipótese do Mercado Eficiente.

Com base em regressões com dados em painéis e vetores autorregressivos (VAR), nossos resultados indicaram que o Google pode ser utilizado como uma ferramenta de previsão de desempenho de volume, retorno e volatilidade das ações e índices da B3, corroborando, por exemplo, com os estudos de Bijl et al. (2016), Da et al. (2011), Vozlyublennaia (2014). Dentre as 11 hipóteses sugeridas e analisadas, apenas a hipótese H2b (relação entre volatilidade dos índices de mercado e evolução semanal/anormal de buscas) não foi validada em nenhuma das relações econométricas propostas.

A partir da análise dos resultados econométricos foi possível ressaltar, por exemplo, os seguintes aspectos:

- (i) A relação entre a atenção anormal na palavra-chave “Ibovespa” e o retorno semanal do índice indicou que a frequência de buscas possui um impacto positivo no retorno semanal dos índices de ação Ibovespa de duas semanas posteriores. Isto sugere que as pesquisas são utilizadas antes de uma decisão de compra por parte dos potenciais acionistas, o que, por sua vez, elevará os preços das ações pertencentes ao índice. Resultados alinhados com a hipótese “*price pressure hypothesis*”, de Barber e Odean (2008).
- (ii) Maior nível de pesquisas em palavras que formam o índice de sentimento FEARS Negativo, podem refletir uma piora no retorno dos principais índices acionários na semana posterior, sustentando a hipótese de que os mercados são afetados pelo *noise trading*, de De Long et al. (1990). Nessa mesma linha, o índice FEARS positivo, com defasagem de duas semanas,

possui um impacto positivo e estatisticamente significativo no retorno semanal dos índices acionários.

- (iii) Maior nível de pesquisas em palavras que formam o índice de sentimento FEARS Negativo, podem refletir, em três semanas posteriores, um aumento na volatilidade do mercado. Tal lógica está em linha com a ideia de que a volatilidade é um indicador do medo e preocupação dos mercados.
- (iv) Um aumento na frequência de pesquisas dos tickers de negociação das empresas reflete uma queda futura no retorno anormal dos títulos, bem como um aumento nas buscas pelo nome da empresa reflete uma queda no retorno semanal dos títulos. Tais observações, são opostas as sugeridas pela hipótese “price pressure hypothesis”.

Para complementar a análise, desenvolvemos uma estratégia de alocação de capitais, com base em inteligência artificial de redes neurais. Os parâmetros de entrada dessa estratégia eram formados pela variação da frequência semanal de buscas da palavra “Ibovespa” defasados em 1 a 5 período e pelos índices de sentimento FEARS positivo e negativo.

Das 1.600 configurações de rede desenvolvidas, 1.401 obtiveram mais acertos do que erros em relação ao resultado esperado (índice vai valorizar ou não na semana) durante a etapa de treinamento, representando cerca de 87,5% de acerto. A rede com mais acertos na fase de treinamento obteve 59,02% de acerto na fase de teste. A topologia de rede utilizada no MATLAB foi bastante eficiente às necessidades do trabalho, podendo dessa forma, ser utilizado com outros tipos de variáveis de entrada.

A partir de duas simulações de investimentos, comparamos as rentabilidades das estratégias baseadas em dados do Google com a rentabilidade do índice Ibovespa. No primeiro exemplo, para um montante de R\$ 10.000 disponíveis por semana, o modelo proposto pela rede neural obteve um retorno acumulado de R\$ 1.655,00, enquanto o índice Ibovespa obteve retorno acumulado de R\$ 1.092,00, resultando em uma diferença de R\$ 563,00. Já no segundo exemplo, para um montante de R\$ 10.000 disponíveis no início do período, o modelo proposto pela rede neural acumulou um total de R\$ 11.662,00, representando um aumento de capital de 17% ou R\$ 1.662,00, enquanto o índice Ibovespa acumulou um total de R\$ 10.913,00, representando um aumento de capital de 9% ou R\$ 913,00.

Concluindo, essa pesquisa colabora para o debate acerca da influência de aspectos relacionados a atenção e ao sentimento dos investidores no mercado financeiro. Apesar das limitações da ferramenta e as variáveis comportamentais serem difíceis de mensurar e muitas vezes abstratas, o trabalho apresentou indícios positivos de que o Google Trends pode ser uma ferramenta de auxílio a tomada de decisões.

8.1 LIMITAÇÕES E SUGESTÕES DE TRABALHOS FUTUROS

As consultas no Google são direcionadas principalmente por investidores individuais que operam valores bem menos significativos que investidores institucionais, uma vez que estes utilizam plataformas profissionais, como o terminal Bloomberg, para consultar dados financeiros de empresas e notícias atualizadas dos acontecimentos do mercado. Dessa forma, operações com volumes financeiros de alto montante e que afetam o desempenho do mercado podem não ser refletidas na evolução do índice de buscas do Google Trends.

Outro aspecto limitante do trabalho é a mensuração do sentimento dos usuários do Google ao buscarem informações sobre tickers e nomes das empresas listadas na B3. Essa dificuldade impede uma análise mais aprofundada de qual direcionamento o mercado está tomando.

Sugerimos para trabalhos futuros a utilização de dados diários do Google Trends para analisar o impacto no mercado brasileiro com uma maior precisão e a construção das análises utilizando a atenção em empresas/tickers em conjunto com dados relacionados ao sentimento FEARS.

REFERÊNCIAS

- ALMEIDA, C. DE P. X. **Aplicação de redes neurais artificiais na previsão do desempenho de ofertas públicas iniciais**. 2013.
- ANDREI, D.; HASLER, M. Investor Attention and Stock Market Volatility. **SSRN Electronic Journal**, 2013.
- AOUADI, A.; AROURI, M.; TEULON, F. Investor attention and stock market activity: Evidence from France. **Economic Modelling**, v. 35, p. 674–681, 2013.
- ASIF KHAN, M. et al. A new sentiment index for the Islamic stock market. **Investment Analysts Journal**, v. 48, n. 2, p. 146–172, 2019.
- AZEVEDO, L. F. P. **Impactos econômicos e financeiros de notícias**. 2017.
- BAKER, M.; WURGLER, J. Investor sentiment and the cross-section of stock returns. **Journal of Finance**, v. 61, n. 4, p. 1645–1680, 2006.
- BAKER, M.; WURGLER, J. Investor sentiment in the stock market. **Journal of Economic Perspectives**, v. 21, n. 2, p. 129–151, 2007.
- BANK, M.; LARCH, M.; PETER, G. Google search volume and its influence on liquidity and returns of German stocks. **Financial Markets and Portfolio Management**, v. 25, n. 3, p. 239–264, 2011.
- BARBER, B. M.; ODEAN, T. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. **Review of Financial Studies**, v. 21, n. 2, p. 785–818, 2008.
- BARBERIS, N.; THALER, R. Survey of Behavioral Finance. **NBER Working Paper Series**, p. 2–67, 2002.
- BARONE, D. **Sociedades artificiais: a nova fronteira da inteligência das máquinas**. Porto Alegre: Bookman, 2003.
- BIJL, L. et al. Google searches and stock returns. **International Review of Financial Analysis**, v. 45, p. 150–156, 2016.
- BOLLEN, J.; MAO, H.; ZENG, X. Twitter mood predicts the stock market. **Journal of Computational Science**, v. 2, n. 1, p. 1–8, 2011.
- BORDINO, I. et al. Web search queries can predict stock market volumes. **PLoS ONE**, v. 7, n. 7, 2012.
- BREALEY, R. A.; MYERS, S. C. **Princípios de finanças empresariais**. 3 ed. ed.

Portugal: 1992.

BROCHADO, A. O. Google Search Based Sentiment indexes. **IIMBManagement Review**, 2019.

BROWN, G. W.; CLIFF, M. T. Investor sentiment and asset valuation. **Journal of Business**, v. 78, n. 2, p. 405–440, 2005.

BRUNI, A. L.; FAMÁ, R. **Eficiência, previsibilidade dos preços e anomalias em mercados de capitais : teoria e evidências**. 1998.

CAMPBELL, J. Y.; LO, A. W.; MACKINLAY, A. C. The Econometrics of Financial Markets. p. 149–180, 1997.

CARAÇA, L. VI. N. **Relação entre o desempenho do mercado brasileiro de ações e o Google Trends**.2019.

CARHART, M. M. On persistence in mutual fund performance. **Journal of Finance**, v. 52, n. 1, p. 57–82, 1997.

CARNEIRO, H. A.; MYLONAKIS, E. Google trends: A web-based tool for real-time surveillance of disease outbreaks. **Clinical Infectious Diseases**, v. 49, n. 10, p. 1557–1564, 2009.

CHALLET, D.; BEL HADJ AYED, A. Predicting Financial Markets with Google Trends and Not so Random Keywords. **SSRN Electronic Journal**, p. 1–9, 2013.

CHAU, M.; VAYANOS, D. The Society for Financial Studies Strong-Form Efficiency with Monopolistic Insiders. **he Review of Financial Studies**, v. 21, n. 5, p. 2275–2306, 2008.

CHEMMANUR, T.; YAN, A. Product market advertising and new equity issues. **Journal of Financial Economics**, v. 92, n. 1, p. 40–65, 2009.

CHEN, S. Google Search Volume : Influence and Indication for the Dutch Stock Market. 2011.

CHOI, H.; VARIAN, H. Predicting the present Choi Varian. **Http://People.Ischool.Berkeley.Edu**, 2011.

CORREIA, L. **Utilização de redes neurais na identificação de tendências de desempenho das Ofertas Públicas Iniciais**. 2015.

CORWIN, S. A.; COUGHENOUR, J. F. Limited attention and the allocation of effort in securities trading. **Journal of Finance**, v. 63, n. 6, p. 3031–3067, 2008.

DA, Z.; ENGELBERG, J.; GAO, P. The sum of all FEARS investor sentiment and asset prices. **Review of Financial Studies**, v. 28, n. 1, p. 1–32, 2015.

- DAMODARAN, A. **Avaliação de investimentos: Ferramentas e Técnicas para a Determinação do valor de qualquer ativo**. Qualitymar ed. Rio de Janeiro: 1997.
- DE LONG, J. B; SHLEIFER, A.; SUMMERS, L. H.; WALDMANN, R. Noise trader risk in financial markets. **Journal of political Economy**, **98(4)**, p. 703–738, 1990.
- DELLAVIGNA, S.; POLLET, J. Investor Inattention, Firm Reaction, and Friday Earnings Announcements. **NBER Working Paper No. 11683**, 2005.
- DIMPFL, T.; JANK, S. Can internet search queries help to predict stock market volatility? **CFR working paper, No. 11-15**, 2011.
- ELTON, E. J.; GRUBER, M. **Modern Portfolio Theory and Investment Analysis**. 5. ed. ed. USA: 1995.
- ENGELBERG, J.; SASSEVILLE, C.; WILLIAMS, J. Market madness? the case of Mad Money. **Management Science**, v. 58, n. 2, p. 351–364, 2012.
- FAMA, E. Efficient Capital Markets: A Review of the Theory. **The Journal of Finance**, v. 25, n. 2, p. 383–417, 1970.
- FAMA, E. F. Efficient Capital Markets: II. **The Journal of Finance**, v. 46, n. 5, p. 1575, 1991.
- FAMA, E. F.; FRENCH, K. R. **Fama-French APT three-factor model** **Journal of Financial Economics**, 1993.
- FANG, L.; PERESS, J. Media Coverage and the Cross-Section of Stock Returns: **The Journal of Finance** **64(5)**, 2009.
- FERNANDES, C. M. DA A. Sentimento e Rendibilidade nos Mercados de Capitais. p. 243, 2015.
- FISCHER BLACK. Noise. **The Journal of Finance**, v. XLI, n. 3, p. 528–543, 1986.
- FISHER, K. L.; STATMAN, M. Investor Sentiment and Stock Returns. **Financial Analysts Journal**, v. 56, n. 2, p. 16–23, 2000.
- GAO, Z.; REN, H.; ZHANG, B. Googling Investor Sentiment around the World. **Journal of Financial and Quantitative Analysis**, v. 55, n. 2, p. 549–580, 2020.
- GROSSMAN, S. J.; STIGLITZ, J. E. On the impossibility of informationally Efficient Markets. **The American economic review**, **70(3)**, p. 393–408, 1980.
- GRULLON, G.; KANATAS, G.; WESTON, J. P. Advertising, breadth of ownership, and liquidity. **Review of Financial Studies**, v. 17, n. 2, p. 439–461, 2004.
- GUZELLA, M. DOS S. Investor Attention in the Brazilian Stock Market : Essays in Behavioral Finance. 2020.

- HALFELD, M.; TORRES, F. DE F. L. Finanças Comportamentais: aplicações no contexto brasileiro. **Rae**, v. 41, n. 2, p. 64–71, 2001.
- Haugen, R. A. **Os Segredos da Bolsa**. São Paulo: 2000, 2000.
- HAYKIN, S. **Redes neurais: princípios e prática**. Porto Alegre: Bookman, 2001.
- HIRSHLEIFER, D.; TEOH, S. H. Limited attention, information disclosure, and financial reporting. **Journal of Accounting and Economics**, 2003.
- HUBERMAN, G.; REGEV, T. Contagious speculation and a cure for cancer: Anonevent that made the stiock market soar. **Journal of Finance**, v. 56, n. 1, p. 387–396, 2001.
- IBGE. **PIB a preços de mercado**. 2021
- JEGADEESH, N. **Evidence of Predictable Behavior of Security Returns****The Journal of Finance**, 1990.
- JENSEN, M. C. Some Anomalous Evidence Regarding Market Efficiency. **Journal of Financial Economics**, v. 6, n. July 2002, p. 95–101, 1978.
- JOSEPH, K.; BABAJIDE WINTOKI, M.; ZHANG, Z. Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. **International Journal of Forecasting**, v. 27, n. 4, p. 1116–1127, 2011.
- JUNIOR, L. F. P. A. **Análise do modelo dos três fatores aplicado à BM&F Bovespa**. [s.l: s.n.].
- KAHNEMAN, D. **Attention and effort**. 1973.
- KAHNEMAN, D.; TVERSKY, A. Prospect theory: An analysis of decision under risk. **Econometrica**, v. 1, p. 143–172, 1979.
- KEYNES, J. M. The General Theory of Employment, Interest and Money. **Journal of the American Statistical Association**, v. 31, n. 196, p. 791, 1936.
- KEZAM MALAGA, F.; SECURATO, J. R. Aplicação do Modelo de Três Fatores de Fama e French no Mercado Acionário Brasileiro – Um Estudo Empírico do Período 1995-2003. p. 1–16, 2003.
- KOSTOPOULOS, D.; MEYER, S.; UHR, C. Google search volume and individual investor trading. **Journal of Financial Markets**, v. 49, p. 100544, 2020.
- KRISTOUFEK, L. Can google trends search queries contribute to risk diversification? **Scientific Reports**, v. 3, p. 1–5, 2013.
- KYLE, A. **Continuous Auctions and Insider Trading****Econometrica**, 1985.
- LATOEIRO, C. J. P. Departamento de Economia Política Pesquisar para decidir : O

- Google como barómetro da atenção do investidor. 2012.
- LEE, CHARLES ; SHLEIFER, ANDREI ; THALER, R. H. Investor sentiment and the closed-end-puzzle. **The Journal of Finance**, **46 (1)**, p. 75–109, 1991.
- LEMMON, M. L. .; PORTNIAGUINA, E. V. Consumer Confidence and Asset Prices: Some Empirical Evidence. **SSRN Electronic Journal**, 2002.
- LO, A. W.; MACKINLAY, A. C. Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test. **NBER Working Paper No. 2168**, p. 17–46, 1987.
- MALKIEL, B. G. American Finance Association Returns from Investing in Equity Mutual Funds 1971 to 1991. **The Journal of Finance**, v. 50, n. 2, p. 549–572, 1995.
- MALKIEL, B. G. The Efficient Market Hypothesis and Its Critics. **Journal of Economic Perspectives** **17(1)**, p. 59–82, 2003.
- MALONEY, M. T.; MULHERIN, J. H. H. The Stock Price Reaction to the Challenger Crash: Information Disclosure in an Efficient Market. 1998.
- MARKOWITZ, H. Portfolio Selection. **The Journal of Finance**, 1952.
- MARTINELLI, E. **Extração de conhecimento de Redes Neurais Artificiais**. 1999.
- MERTON, R. C. A Simple Model of Capital Market Equilibrium with Incomplete Information. **The Journal of Finance**, v. 42, n. 3, p. 483–510, 1987.
- MIRAGAYA, F. C. S. DA C. **O Efeito dos Retornos Anormais nas Buscas por Informação dos Investidores**. 2017.
- MOAT, H. S. et al. Quantifying wikipedia usage patterns before stock market moves. **Scientific Reports**, v. 3, p. 1–5, 2013.
- MONDRIA, J.; WU, T. Asymmetric Attention and Stock Returns. **SSRN Electronic Journal**, 2012.
- MONDRIA, J.; WU, T.; ZHANG, Y. The determinants of international investment and attention allocation: Using internet search query data. **Journal of International Economics**, v. 82, n. 1, p. 85–95, 2010.
- NETO, J. N. DE M. **Fractais e redes neurais artificiais aplicados à previsão de retorno de ativos financeiros brasileiros**. 2014.
- ODEAN, T. Do Investors Trade Too Much? **The American Economic Review**, **89 (5)**, p. 1279–1298, 1999.
- PASHLER, H. E. **The Psychology of Attention**. 1998.
- PENG, L. et al. Investor Attention, Overconfidence and Category Learning. **NBER**

Working Paper No. 11400, 2005.

PERLIN, M. S. et al. Can we predict the financial markets based on google's search queries? **Journal of Forecasting**, v. 36, n. 4, p. 454–467, 2016.

PREIS, T.; MOAT, H. S.; EUGENE STANLEY, H. Quantifying trading behavior in financial markets using google trends. **Scientific Reports**, v. 3, p. 1–6, 2013.

PREIS, T.; REITH, D.; STANLEY, H. E. Complex dynamics of our economic life on different scales: Insights from search engine query data. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 368, n. 1933, p. 5707–5719, 2010.

RAMOS, H. P.; RIBEIRO, K. K. M.; PERLIN, M. S. the Forecasting Power of Internet Search Queries in the Brazilian Financial Market. **RAM. Revista de Administração Mackenzie**, v. 18, n. 2, p. 184–210, 2017.

RENAULT, T. Intraday online investor sentiment and return patterns in the U.S. stock market. **Journal of Banking and Finance**, v. 84, n. 0, p. 25–40, 2017.

RICH, E.; KNIGHT, K. **Inteligência Artificial**. [s.l.] Makron Books, 1993.

ROBERTS, H. **Statistical versus clinical prediction of the stock market**. 1967.

RODOLFO, E. R.; BARBEDO, C. H. DA S.; VAL, F. DE F. Relevance of the Attention Effect in the Brazilian Stock Market. **Latin American Business Review**, v. 18, n. 2, p. 165–187, 2017.

ROSS, S.; WESTWFIELD, R.; JAFFE, J. **Administração Financeira**. São Paulo: 1995, 1995.

SCBIEHLL, E. O Efeito da Divulgação das Demonstrações Financeiras no mercado de Capitais Brasileiro: Um Estudo sobre a Variação no Preço das Ações. v. 655, p. 114, 1996.

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

SCHMIDT, T.; VOSEN, S. Forecasting Private Consumption: Survey-based Indicators vs. Google Trends. **Ruhr Economic Papers, No. 155**, 2009.

SEASHOLES, M. S.; WU, G. Predictable behavior, profits, and attention. **Journal of Empirical Finance**, v. 14, n. 5, p. 590–610, 2007.

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

SHLEIFER, A.; VISHNY, R. W. The limits of arbitrage. **The Journal of Finance**, v.

52, p. 79–101, 1997.

SHLEIFER, A.; VISHNY, R. W. The limits of arbitrage. **Advances in Behavioral Finance**, v. 2, n. 1, p. 79–101, 2005.

SIGANOS, A.; VAGENAS-NANOS, E.; VERWIJMEREN, P. Divergence of sentiment and stock market trading. **Journal of Banking and Finance**, v. 78, n. March, p. 130–141, 2017.

SIMON, H. A. **Theories of Bounded Rationality Decision and Organization**, 1972.

STREICHER, A. R. Stock Returns and Google Search Volume Data – An analysis on the Portuguese and American market –. v. D, 2019.

TAKEDA, F.; WAKAO, T. Google search intensity and its relationship with returns and trading volume of Japanese stocks. **Pacific Basin Finance Journal**, v. 27, n. 1, p. 1–18, 2014.

TANTAOPAS, P.; PADUNGSAKSAWASDI, C.; TREEPONGKARUNA, S. Attention effect via internet search intensity in Asia-Pacific stock markets. **Pacific Basin Finance Journal**, v. 38, p. 107–124, 2016.

TETLOCK, P. C. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. v. LXII, n. 3, p. 1139–1168, 2007.

THALER, R. Behavioral Economics: Past, present and future. **American Economic Review**, v. 106, n. 7, p. 1577–1600, 2016.

VAN NIEUWERBURGH, S.; VELDKAMP, L. Information acquisition and under-diversification. **Review of Economic Studies**, v. 77, n. 2, p. 779–805, 2010.

VIANA, L. F. **Google Trends e o comportamento do mercado acionário brasileiro**. 2017.

VLASTAKIS, N.; MARKELLOS, R. N. Information demand and stock market volatility. **Journal of Banking and Finance**, v. 36, n. 6, p. 1808–1821, 2012.

VOZLYUBLENNAIA, N. Investor attention, index performance, and return predictability. **Journal of Banking and Finance**, v. 41, n. 1, p. 17–35, 2014.

WHALEY ROBERT. The investor fear gauge. **The journal of portfolio Management**, n. Vix, 2000.

YOSHINAGA, C. E. **A relação entre índice de sentimento de mercado e as taxas de retorno das ações: Uma análise com dados em painel**. 2009.

YOSHINAGA, C.; ROCCO, F. Atenção do Investidor: O Volume de Buscas no Google é Capaz de Prever os Retornos de Ações? v. d, 2019.

ZHANG, C. Defining , Modeling , and Measuring Investor Sentiment. **Working Paper**, n. April, p. 0–42, 2008.

ZHI DA, JOSEPH ENGELBEG, AND P. G. In Search of Attention ZHI DA, JOSEPH ENGELBERG, and PENGJIE GA (). **Journal of Finance**, v. LXVI, n. 5, p. 1461–1499, 2011.

ZWEIG, M. E. An Investor Expectations Stock Price Predictive Model Using Closed-End Fund Premiums. **The Journal of Finance**, Vol. 28, No. 1 (Mar., 1973), v. 73, p. 67–78, 1973.

APÊNDICE A - PALAVRAS DE SENTIMENTO DE DA ET. AL (2015)

| PALAVRAS INGLÊS | TRADUÇÃO PARA PORTUGUÊS |
|-----------------------|---------------------------|
| 401k | *Não Traduzida |
| 401k Contribution | *Não Traduzida |
| American Savings | Poupança BRASILEIRA |
| Bankrupt | Falido |
| Bankruptcy | Falência |
| Bankruptcy Attorney | Advogado De Falências |
| Bankruptcy Chapter 7 | *Não Traduzida |
| Bankruptcy Court | Tribunal De Falências |
| Bankruptcy Law | Lei De Falências |
| Bankruptcy Laws | Leis De Falência |
| Bargain Trader | Comerciante De Pechinchas |
| Benefit | Benefício |
| Benefits | Benefícios |
| Budget Deficit | Déficit Orçamentário |
| Business Partnership | Parceria Comercial |
| Buy Gold | Comprar Ouro |
| Capitalization | Capitalização |
| Car Donate | Doação De Carro |
| Chapter 7 | *NÃO TRADUZIDA |
| Chapter 13 | *NÃO TRADUZIDA |
| Chapter 13 Bankruptcy | *Não Traduzida |
| Charity | Caridade |
| College Savings | Poupança Para Faculdade |
| Community Bank | Banco Comunitário |
| Contribution Limits | Limites De Contribuição |
| Cooperative Bank | Banco Cooperativo |
| Cost Accounting | Contabilidade De Custos |
| Cost Of Life | Custo De Vida |
| Creditor | Credor |
| Crisis | Crise |
| Default | Padrão |
| Deferred Compensation | Compensação Diferida |
| Deficit | Déficit |
| Depression | Depressão |
| Donation | Doação |

| Palavras Inglês | Tradução |
|-------------------------|---|
| Economy | Economia |
| Entrepreneur | Empresário |
| Entrepreneurial | Empresarial |
| Entrepreneurship | Empreendedorismo |
| Equity | Patrimônio |
| Equity Bank | Banco De Patrimônio Líquido |
| Equity Fund | Fundo De Patrimônio |
| Equity Line | Linha De Equidade |
| Equity Loan | Empréstimo De Capital De Giro |
| Expense | Despesa |
| Expenses | Despesas |
| Federal Poverty Level | Nível De Pobreza Nacional |
| File Bankruptcy | Declarar Falência |
| Filing Bankruptcy | Abrir Falência |
| Financial Crisis | Crise Financeira |
| For Profit | Para Lucro |
| Frugal | Frugal |
| Gdp | Pib |
| Gold | Ouro |
| Gold Price | Preço Ouro |
| Gold Prices | Preços Ouro |
| Great Depression | Grande Depressão |
| Health Insurance | Seguro Saúde |
| Home Equity | Patrimônio Familiar |
| Home Equity Line | *Não Traduzida |
| Home Equity Loan | Empréstimo Para Aquisição De Casa Própria |
| Housing Allowance | Subsídio De Moradia |
| Inflation | Inflação |
| Inflation Calculator | Calculadora De Inflação |
| Inflation Rate | Taxa De Inflação |
| INFLATION Rates | Taxas De Inflação |
| Ira | *Não Traduzida |
| Ira CONTRIBUTION | *NÃO TRADUZIDA |
| Ira Contribution Limits | *Não Traduzida |
| Ira Limits | *Não Traduzida |
| Lay Off | Demitido |
| Limited Partnership | Parceria Limitada |
| Liquidation | Liquidação |
| Margin | Margem |
| Net Worth | Patrimônio Líquido |
| Non-Profit Jobs | Empregos Sem Fins Lucrativos |
| Poor Credit | Crédito Baixo |
| Poverty | Pobreza |

| Palavras Inglês | Tradução |
|--------------------------------|-------------------------|
| POVERTY Level | Nível De POBREZA |
| POVERTY Line | LINHA DA POBREZA |
| POVERTY Rate | Taxa De POBREZA |
| POVERTY Statistics | Estatísticas De POBREZA |
| PRICE Of Gold | PREÇO Do Ouro |
| PRIVATE Equity | PRIVATE Equity |
| Profit | Lucro |
| Profit Margin | Margem De Lucro |
| Profitable | Rentável |
| Recession | Recessão |
| Rich | Riqueza |
| Roth Contribution | *Não Traduzida |
| Roth Ira | *Não Traduzida |
| ROTH Ira CONTRIBUTION | *NÃO TRADUZIDA |
| Savings Account | Conta Poupança |
| Savings | Poupança |
| SAVINGS Bond | Título De Renda Fixa |
| SAVINGS Bonds | Títulos De Renda Fixa |
| SAVINGS Calculator | CALCULADORA DE POUPANÇA |
| SOCIAL Security | *NÃO TRADUZIDA |
| SOCIAL Security Benefit | *NÃO TRADUZIDA |
| SOCIAL Security Card | *NÃO TRADUZIDA |
| SOCIAL Security Office | *NÃO TRADUZIDA |
| Successful Business | Negócios De Sucesso |
| Tariff | Tarifa |
| Tariffs | Tarifas |
| Taxes | Impostos |
| The CRISIS | A Crise |
| The Deficit | O Déficit |
| The Depression | A Depressão |
| The Great Depression | A Grande Depressão |
| Thrift Savings | *Não Traduzida |
| TRADE Deficit | Déficit Comercial |
| Unemployed | Desempregado |
| Unemployment | Desemprego |
| Us DEFICIT | Déficit Brasileiro |
| US Inflation | INFLAÇÃO BRASILEIRA |
| Us Poverty | Pobreza Brasileira |
| WORKERS COMPENSATION Insurance | Seguro Desemprego |
| WORLD Poverty | Pobreza MUNDIAL |

FONTE: O autor (2021) com base em Da et. al (2015)

APÊNDICE B - SELEÇÃO DE PALAVRAS DE SENTIMENTOS

| Palavra | Palavra | Palavra | Palavra |
|-------------------------|-------------------------------|---------------------|-----------------------|
| A Crise | Declarar Falência | Fundo De Patrimônio | Pobreza |
| Abrir Falência | Déficit | Grande Depressão | Pobreza Brasileira |
| Aluguel | Déficit Brasileiro | Hipoteca | Pobreza Mundial |
| Aposentadoria | Déficit Comercial | Impostos | Poupança |
| Auxilio | Déficit Orçamentário | Inflação | Poupança Brasileira |
| Banco | Demitido | Inflação Brasileira | Preço Do Ouro |
| Banco Comunitário | Depressão | Informalidade | Private Equity |
| Banco Cooperativo | Desconto | INSS | Recessão |
| Benefício | Desempregado | Investir Em Ações | Renda Fixa |
| Benefícios | Desemprego | IPO | Rentável |
| Bolsa Família | Desigualdade | Juros | Riqueza |
| Calculadora De Inflação | Despesa | Lei De Falências | Salário Mínimo |
| Calculadora De Poupança | Despesas | Linha Da Pobreza | Seguro Desemprego |
| Capitalização | Divida | Liquidação | Seguro Saúde |
| Caridade | Dividas | Lucro | Selic |
| CDI | Doação | Margem | Serasa |
| Cesta Básica | Dólar | Margem De Lucro | Subsídio |
| Comprar Ouro | Economia | Negativo No Banco | Subsídio De Moradia |
| Conta Atrasada | Empreendedorismo | Negócios De Sucesso | Tarifa |
| Conta Poupança | Empresarial | O Déficit | Tarifas |
| Contabilidade De Custos | Empresário | Orçamento Familiar | Taxa De Corretagem |
| Crédito Baixo | Empréstimo | Ouro | Taxa De Financiamento |
| Credor | Empréstimo De Capital De Giro | Parceria Comercial | Taxa De Inflação |
| Crise | Falência | Patrimônio | Taxa De Pobreza |
| Crise Financeira | Falido | Patrimônio Líquido | Taxas De Inflação |
| Custo De Vida | FGTS | PIB | Tesouro Direto |
| Debito | Financiamento | Plano De Saúde | |

FONTE: O autor (2021)