

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO SOCIOECONÔMICO
CURSO DE GRADUAÇÃO EM CIÊNCIAS ECONÔMICAS

Bruno de Mattos Ritter (20101195)

Previsão de volatilidade em ações brasileiras: o papel preditivo do Google Trends

Florianópolis
2026

Bruno de Mattos Ritter (20101195)

Previsão de volatilidade em ações brasileiras: o papel preditivo do Google Trends

Trabalho de Conclusão de Curso submetido ao Curso de Graduação em Ciências Econômicas do Centro Socioeconômico da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Bacharel em Ciências Econômicas.
Orientador: Prof. Guilherme Valle Moura, Dr.

Florianópolis
2026

Ficha catalográfica gerada por meio de sistema automatizado gerenciado pela BU/UFSC.
Dados inseridos pelo próprio autor.

Ritter, Bruno de Mattos
Previsão de volatilidade em ações brasileiras: o papel preditivo do Google Trends / Bruno de Mattos Ritter ; orientador, Guilherme Valle Moura, 2026.
40 p.

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Santa Catarina, Centro Socioeconômico, Graduação em Ciências Econômicas, Florianópolis, 2026.

Inclui referências.

1. Ciências Econômicas. 2. Volatilidade realizada. 3. Google Trends. 4. Atenção do investidor. 5. Modelo eGARCH. I. Moura, Guilherme Valle. II. Universidade Federal de Santa Catarina. Graduação em Ciências Econômicas. III. Título.

Bruno de Mattos Ritter (20101195)

Previsão de volatilidade em ações brasileiras: o papel preditivo do Google Trends

Florianópolis, 06/07/2026

O presente Trabalho de Conclusão de Curso foi avaliado e aprovado pela banca examinadora composta pelos seguintes membros:

Prof. Guilherme Valle Moura, Dr.
Universidade Federal de Santa Catarina

Prof. Pedro Chaim, Dr.
Universidade Federal de Santa Catarina

Prof. João Caldeira, Dr.
Universidade Federal de Santa Catarina

Certifico que esta é a **versão original e final** do Trabalho de Conclusão de Curso que foi julgado adequado para obtenção do título de Bacharel em Economia por mim e pelos demais membros da banca examinadora.

ICP-Edu
Documento assinado digitalmente
Guilherme Valle Moura
Data: 07/07/2026 16:25:23-0300
CPF: ***.234.686-**
Verifique as assinaturas em https://v.ufsc.br

Prof. Guilherme Valle Moura, Dr.
Orientador

Florianópolis, 2026.

Resumo

Este trabalho investiga se o *Search Volume Index* (SVI) do Google Trends, utilizado como *proxy* de atenção do investidor, possui poder preditivo sobre a volatilidade realizada semanal de ações brasileiras listadas na B3. A pesquisa compara três modelos de previsão: o passeio aleatório (M0), como referência ingênua; o eGARCH(1,1) (M1); e o eGARCH-X(1,1) (M2), que incorpora a atenção anormal de buscas. A variável exógena é o desvio logarítmico do SVI em relação à sua mediana das oito semanas anteriores. A comparação baseia-se em um exercício de previsão fora da amostra, com janela móvel de 156 semanas e previsões de um passo à frente para um painel de 260 ações, avaliado pela função de perda QLIKE combinada ao procedimento *Model Confidence Set* (MCS). Os resultados mostram que ambos os modelos eGARCH superam amplamente o passeio aleatório, mas o conjunto de confiança do MCS, ao nível de 10%, contém exclusivamente o eGARCH puro. Portanto, a inclusão da atenção de buscas não melhora a previsão da volatilidade no agregado, resultado robusto a diferentes critérios de filtragem das séries. Ao mesmo tempo, o coeficiente associado ao SVI é estatisticamente significativo em pelo menos metade das janelas para 101 dos 260 ativos. Além disso, os tickers com SVI significativo em mais de 70% da previsão fora da amostra possuem maior volatilidade realizada média que os demais. Conclui-se que a significância estatística do coeficiente de atenção e a sua utilidade preditiva fora da amostra são fenômenos empiricamente independentes, e que a atenção anormal do investidor segue a volatilidade realizada, em vez de antecedê-la. Os achados contribuem para a literatura de finanças comportamentais e atenção do investidor em mercados emergentes, sinalizando que a atenção captada por buscas na internet tem relevância estatística para a dinâmica da volatilidade, mas benefício preditivo incremental limitado e heterogêneo entre os ativos.

Palavras-chave: Volatilidade realizada; Google Trends; Atenção do investidor; Modelo eGARCH; *Model Confidence Set*.

Abstract

This study investigates whether the Google Trends *Search Volume Index* (SVI), used as a proxy for investor attention, has predictive power over the weekly realized volatility of Brazilian stocks listed on B3. The analysis compares three forecasting models: a random walk benchmark (M0), serving as a naïve reference model; an eGARCH(1,1) model (M1); and an eGARCH-X(1,1) model (M2), which incorporates abnormal search attention. The exogenous variable is defined as the logarithmic deviation of the SVI from its median over the previous eight weeks. Model performance is evaluated through an out-of-sample forecasting exercise using a rolling window of 156 weeks and one-step-ahead forecasts for a panel of 260 stocks. Forecast accuracy is assessed using the QLIKE loss function combined with the *Model Confidence Set* (MCS) procedure. The results show that both eGARCH specifications substantially outperform the random walk benchmark. However, at the 10% significance level, the MCS confidence set contains only the standard eGARCH model. Therefore, the inclusion of search-based attention does not improve volatility forecasts in the aggregate, a result that remains robust across different data-filtering criteria. At the same time, the coefficient associated with the SVI is statistically significant in at least half of the estimation windows for 101 of the 260 assets. Furthermore, stocks for which the SVI coefficient is significant in more than 70% of the out-of-sample forecasting period exhibit higher average realized volatility than the remaining stocks. These findings indicate that the statistical significance of the attention coefficient and its out-of-sample predictive usefulness are empirically distinct phenomena. The evidence suggests that abnormal investor attention tends to follow realized volatility rather than precede it. The study contributes to the literature on behavioral finance and investor attention in emerging markets by showing that internet search activity carries statistical relevance for volatility dynamics, while providing limited and heterogeneous incremental forecasting benefits across assets.

Keywords: Realized volatility; Google Trends; Investor attention; eGARCH model; *Model Confidence Set*.

Sumário

1 INTRODUÇÃO 7
2 REFERENCIAL TEÓRICO 8
2.1 Eficiência de mercado e o pressuposto de atenção plena . . . 9
2.2 Atenção limitada e finanças comportamentais 9
2.3 Atenção do investidor e a dinâmica de preços e volatilidade . 10
2.4 A mensuração da atenção: de proxies indiretas ao volume de buscas 11
2.5 Proxies alternativas de atenção: redes sociais e plataformas de vídeo 12
2.6 Posicionamento e contribuição deste trabalho 12
3 DADOS E MÉTODO 13
3.1 Dados 13
3.1.1 Coleta do Search Volume Index via Google Trends 13
3.1.2 Universo de ativos e critérios de elegibilidade 14
3.2 Metodologia 15
3.2.1 Volatilidade realizada 15
3.2.2 A variável de atenção: construção e estacionariedade 16
3.2.3 Exercício de previsão fora da amostra com janela móvel . . 17
3.2.4 Modelo 0: passeio aleatório 18
3.2.5 Modelo 1: eGARCH 18
3.2.6 Modelo 2: eGARCH-X 20
3.2.7 Identificação de γ e tratamento das séries de SVI 20
3.2.8 Função de perda QLIKE 21
3.2.9 Model Confidence Set 21
4 RESULTADOS 22
4.1 Composição do painel 23
4.2 Acurácia preditiva: QLIKE e Model Confidence Set 23
4.3 Diagnóstico dos resíduos 24
4.4 Robustez à filtragem das séries de SVI 24
4.5 Sobre os ativos com SVI significativo 25
5 CONCLUSÃO 27

1 APÊNDICE A - Resultados das estimações por ticker 29
REFERÊNCIAS 37

1 INTRODUÇÃO

O comportamento dos investidores reflete-se diretamente na dinâmica da volatilidade do mercado financeiro. Em um mercado eficiente, o preço dos ativos incorpora instantaneamente toda a informação disponível (Fama, 1970). Logo, eles seguem um passeio aleatório, o que torna o preço futuro imprevisível. Na prática, a atenção que os agentes dedicam a cada ativo varia ao longo do tempo devido à atenção limitada dos agentes (Kahneman, 1973), e essa variação pode afetar o processo de formação de preços de maneiras que modelos tradicionais nem sempre capturam. No Brasil, a base de investidores pessoa física na bolsa de valores Brasil, Bolsa, Balcão (B3) cresceu de forma expressiva na última década (B3 S.A. – Brasil, Bolsa, Balcão, 2025), ampliando o papel que o interesse e a atenção do investidor de varejo exercem sobre a dinâmica dos preços das ações listadas na bolsa.

Nesse contexto, a disseminação do uso de mecanismos de busca na internet abriu uma nova janela para observar, de forma indireta, o comportamento coletivo dos investidores. O Google, como principal plataforma de busca utilizada no Brasil (StatCounter Global Stats, 2025), registra o interesse das pessoas por determinados temas ao longo do tempo por meio do *Search Volume Index* (SVI), disponibilizado pela ferramenta Google Trends. A ideia central é que o ato de buscar por um ativo antes de negociá-lo deixa um rastro digital mensurável: um aumento no volume de buscas pode sinalizar maior atenção do mercado e, potencialmente, antecipar movimentos de volatilidade.

Perlin et al. (2017) investigaram a interação entre consultas de busca no Google e diferentes aspectos dos mercados acionários em quatro países de língua inglesa, utilizando um modelo vetorial autorregressivo e testes de causalidade de Granger. Os autores identificaram que o aumento na frequência de buscas por termos financeiros pode prever um aumento de volatilidade e queda nos retornos dos índices de mercado na semana seguinte, com o termo *stock* apresentando resultados robustos em todos os países analisados. Essa evidência sugere que os dados de busca capturaram, de forma sistemática, a atenção dos investidores antes de decisões de venda. DA, ENGELBERG e GAO (2011) já haviam apontado na mesma direção, mostrando que aumentos no SVI estão associados a retornos positivos de curto prazo e a maiores retornos no primeiro dia de negociação de IPOs, o que é consistente com a hipótese de que o SVI captura a atenção do investidor de varejo.

Este trabalho estende a utilização do Google Trends ao contexto do mercado acionário brasileiro. Alinha-se a DA, ENGELBERG e GAO (2011) ao adotar o próprio código de negociação (ticker) de cada ação como termo de busca. Isso ocorre porque essa escolha reduz o ruído proveniente de buscas realizadas por consumidores ou por pessoas interessadas em notícias gerais sobre a companhia. Ao pesquisar pelo ticker de uma empresa, os resultados retornados são, em sua maioria, cotações, análises de investimento e conteúdo voltado ao mercado de capitais, tornando o SVI por ticker uma proxy mais precisa e direta da atenção do investidor em relação à ação.

O objetivo central é avaliar se a atenção do investidor, captada semanalmente pelo SVI do Google Trends para os tickers de ações listadas na B3, possui poder preditivo sobre a volatilidade realizada dessas ações em um exercício fora da amostra. Para isso, são estimados e comparados três modelos competidores: um passeio aleatório como referência ingênua; um eGARCH(1,1) como *benchmark* de séries temporais; e um eGARCH-X(1,1) que incorpora a atenção anormal de buscas através do desvio logarítmico do SVI em relação à sua mediana das oito semanas anteriores como variável exógena na equação da variância condicional. A avaliação da acurácia preditiva utiliza a função de perda QLIKE combinada ao procedimento *Model Confidence Set* (MCS), que permite comparar conjuntos de modelos concorrentes sem as limitações do teste de Diebold-Mariano tradicional. Os modelos são estimados com janela móvel de 156 semanas sobre um painel de ações elegíveis da B3, e a base final cobre 260 semanas, da semana encerrada em 02 de abril de 2021 à semana encerrada em 20 de março de 2026.

A análise contribui para o debate sobre o papel da atenção do investidor na dinâmica do mercado de capitais brasileiro, oferecendo evidências baseadas em um exercício preditivo rigoroso, em um universo amplo de ativos e em uma metodologia que concilia a literatura de finanças comportamentais com ferramentas econométricas de previsão de volatilidade.

O restante do estudo está estruturado da seguinte forma: a Seção 2 apresenta o referencial teórico que fundamenta a pesquisa, articulando a hipótese de mercados eficientes e a literatura de atenção limitada; a Seção 3 descreve os dados e detalha a metodologia adotada; a Seção 4 expõe e discute os resultados obtidos com os modelos econométricos propostos; e a Seção 5 apresenta as conclusões finais do estudo.

2 REFERENCIAL TEÓRICO

A investigação sobre o poder preditivo da atenção do investidor situa-se na interseção de duas tradições da teoria de finanças. De um lado, a hipótese dos mercados eficientes, que postula a incorporação instantânea da informação aos preços e a imprevisibilidade dos retornos (Fama, 1970). De outro, a literatura de finanças comportamentais, que reconhece a atenção como um recurso cognitivo escasso e investiga como suas variações afetam a formação de preços e a dinâmica da volatilidade (Kahneman, 1973; DA; ENGELBERG; GAO, 2011). Esta seção percorre esse arcabouço em seis passos: parte do pressuposto de eficiência (Seção 2.1); introduz o conceito de atenção limitada (Seção 2.2); examina o vínculo teórico entre atenção do investidor, preços e volatilidade (Seção 2.3); discute o problema da mensuração da atenção e a emergência do volume de buscas como proxy direta (Seção 2.4); apresenta proxies alternativas extraídas de redes sociais e plataformas de vídeo (Seção 2.5); e, por fim, posiciona a contribuição deste trabalho em relação à literatura (Seção 2.6).

2.1 Eficiência de mercado e o pressuposto de atenção plena

O ponto de partida da discussão é a hipótese dos mercados eficientes (HME), formalizada por Fama (1970). Em um mercado eficiente, os preços dos ativos refletem plenamente toda a informação disponível, de modo que nenhum agente seria capaz de obter retornos anormais de forma sistemática a partir de um dado conjunto informacional. Fama (1970) organiza a hipótese em três formas, conforme o conjunto de informação considerado: a forma fraca, em que os preços incorporam o histórico de preços passados; a forma semiforte, em que incorporam toda a informação publicamente disponível; e a forma forte, em que incorporam inclusive informação privada. Uma implicação central da hipótese, em sua forma fraca, é que os preços seguem um passeio aleatório, tornando o preço futuro imprevisível a partir do passado (Gujarati; Porter, 2011). Trata-se de uma premissa que justifica o uso do passeio aleatório como referência ingênua (Modelo 0) no exercício preditivo deste trabalho.

A HME repousa, contudo, sobre um pressuposto frequentemente implícito: para que a informação seja precificada no instante em que chega, é necessário que os investidores aloquem atenção suficiente ao ativo (DA; ENGELBERG; GAO, 2011). Esse pressuposto é o elo que conecta a teoria da eficiência à literatura de atenção. Perlin et al. (2017) reconhecem que a questão de como a informação atinge os participantes e afeta suas decisões de negociação constitui um dos pilares que originaram a própria teoria da eficiência de mercado (Fama, 1965, 1970). Esses autores destacam ainda que o comportamento individual do investidor é, para a pesquisa, uma espécie de caixa-preta: observam-se apenas os resultados agregados da interação entre os agentes, na forma de preços e volumes negociados, mas nunca a disposição individual que os origina. É justamente essa lacuna — a impossibilidade de observar diretamente o que move cada investidor — que motiva a busca por medidas indiretas e diretas da atenção.

2.2 Atenção limitada e finanças comportamentais

O conceito de atenção limitada tem origem na psicologia cognitiva. Kahneman (1973) caracteriza a atenção como um recurso escasso e finito, que precisa ser alocado seletivamente diante de um volume de informação que excede a capacidade de processamento do indivíduo. Essa ideia é sintetizada na epígrafe de Herbert Simon que abre o trabalho de DA, ENGELBERG e GAO (2011): “a riqueza de informação cria uma pobreza de atenção”. Aplicada aos mercados financeiros, a noção implica que, se a atenção é limitada, a informação não é necessariamente incorporada aos preços de forma instantânea, contrariando o pressuposto da HME e abrindo espaço para que variações na atenção afetem a formação de preços.

A literatura teórica que incorpora a atenção limitada ao apreçamento de ativos é ampla. Merton (1987) propõe um modelo de equilíbrio de mercado com informação

incompleta, no qual os investidores apenas negociam ativos de que têm conhecimento, de modo que o grau de reconhecimento de um ativo afeta seu preço de equilíbrio. Hirshleifer e Teoh (2003) modelam os efeitos da atenção limitada sobre a divulgação de informações e o reporte financeiro, mostrando que a forma de apresentação da informação altera sua incorporação aos preços quando os investidores têm capacidade de processamento restrita. Peng e Xiong (2006) demonstram que a atenção limitada induz os investidores a um aprendizado por categorias, em que processam informação ao nível do mercado ou do setor em detrimento de informação específica da firma. Sims (2003), por sua vez, formaliza a ideia de inatenção racional, segundo a qual os agentes alocam sua capacidade limitada de processamento de informação de maneira ótima. Em conjunto, esses trabalhos fornecem o fundamento teórico de que a atenção é uma variável relevante para a dinâmica do mercado, e não um detalhe comportamental secundário.

2.3 Atenção do investidor e a dinâmica de preços e volatilidade

Estabelecido que a atenção é limitada e relevante, a literatura especifica os canais pelos quais ela afeta o mercado. O canal mais influente é a hipótese da pressão de preços induzida por atenção, proposta por Barber e Odean (2008). O investidor de varejo precisa selecionar entre um conjunto muito amplo de alternativas, e tende a considerar apenas aquelas que capturaram sua atenção. Dentro dessa amostra de possibilidades, o investidor escolhe de acordo com suas preferências individuais. Consequentemente, os investidores individuais são compradores líquidos de ativos em dias de alta volatilidade e quando estão no noticiário. Em contrapartida, os autores concluem que investidores institucionais não demonstram comprar ativos por chamarem atenção, pois dispõem de outros mecanismos para fazer a seleção.

DA, ENGELBERG e GAO (2011) testam empiricamente essa hipótese utilizando o volume de buscas no Google como medida de atenção do investidor de varejo. Os autores constroem a variável de atenção anormal ($ASVI$), definida como o logaritmo do volume de buscas da semana corrente subtraído da mediana das oito semanas anteriores, e encontram que um aumento na atenção está associado a um desempenho superior a 30 pontos-base nas duas semanas seguintes, seguido de reversão quase completa ao longo do ano. O efeito é mais forte entre ações mais negociadas por investidores individuais, e os autores exploram as ofertas públicas iniciais (IPOs) como cenário natural de teste, dado que o volume de buscas existe antes mesmo de o ativo passar a ser negociado.

Embora DA, ENGELBERG e GAO (2011) concentrem-se sobre o retorno, a literatura também vincula a atenção diretamente à volatilidade. Andrei e Hasler (2015) demonstram, em arcabouço teórico e empírico, que a atenção do investidor e a incerteza sobre o aprendizado afetam tanto a volatilidade dos retornos quanto o prêmio de risco, de modo que períodos de maior atenção tendem a coincidir com maior volatilidade. É nessa direção que se inscreve o resultado central de Perlin et al. (2017). Os autores analisaram

quatro mercados acionários de língua inglesa por meio de um modelo vetorial autorregressivo (VAR) com testes de causalidade de Granger, e encontram que um aumento na frequência de buscas pela palavra *stock* prevê um aumento da volatilidade e uma queda dos retornos do índice na semana seguinte. A interpretação proposta é que os investidores realizam buscas antes de decisões sistemáticas de venda, e que o efeito é especialmente forte durante a crise financeira de 2009. Esse achado fundamenta a hipótese central deste trabalho: a de que a variação no volume de buscas carrega informação preditiva sobre a volatilidade futura dos ativos.

2.4 A mensuração da atenção: de proxies indiretas ao volume de buscas

O principal obstáculo para testar as teorias de atenção é que a atenção do investidor não é diretamente observável (DA; ENGELBERG; GAO, 2011). Diante dessa limitação, a literatura recorreu, por muito tempo, a proxies indiretas. Entre elas, destacam-se os retornos extremos, o volume de negociação ou turnover, a presença em notícias e manchetes, os gastos com publicidade e os limites de preço (Barber; Odean, 2008; Gervais; Kaniel; Mingelgrin, 2001). Todas essas medidas assumem como pressuposto que se um ativo apresentou retorno extremo, alto giro ou foi mencionado na mídia, então os investidores teriam prestado atenção a ele. DA, ENGELBERG e GAO (2011) apontam a fragilidade desse pressuposto, uma vez que retorno e giro podem ser determinados por fatores não relacionados à atenção, e a publicação de uma notícia não garante atenção a menos que os investidores efetivamente a leiam.

A contribuição metodológica de DA, ENGELBERG e GAO (2011) é propor o volume de buscas no Google, divulgado pela ferramenta Google Trends sob a forma do *Search Volume Index* (SVI), como uma medida direta e revelada de atenção. Afinal, quem busca por um ativo em um mecanismo de busca está, inequivocamente, prestando atenção nele. Além de direta, a medida é oportuna, pois reflete o interesse no momento em que ocorre, e os autores apresentam evidências de que o SVI capta sobretudo a atenção do investidor individual, ou de varejo, menos sofisticado. Para capturar a atenção dirigida a uma ação específica, DA, ENGELBERG e GAO (2011) utilizam o código de negociação (ticker) como termo de busca (por exemplo, "AAPL" para a Apple). Essa estratégia isola a atenção de quem efetivamente se interessa pelo ativo enquanto investimento, reduzindo a contaminação por buscas de consumidores ou de leitores de notícias gerais sobre a empresa.

A escolha do termo de busca constitui uma decisão metodológica central nessa literatura, e nela os artigos-base divergem. Enquanto DA, ENGELBERG e GAO (2011) utilizam o ticker, Perlin et al. (2017) optam por um conjunto de palavras relacionadas a finanças, selecionadas a partir de um dicionário financeiro e de livros-texto da área. Eles criticam explicitamente o uso restrito de tickers, por considerá-lo um conjunto de expressões demasiado estreito. Outros trabalhos exploram variações dessa decisão: Bordino et al. (2012) mostram que buscas por tickers preveem o volume negociado de ações norte-

americanas; Preis, Moat e Stanley (2013) constroem estratégias de negociação a partir do volume de buscas por termos financeiros; e Vozlyublennaia (2014) examina a relação entre atenção do investidor, desempenho de índices e previsibilidade de retornos. O presente trabalho alinha-se a DA, ENGELBERG e GAO (2011) ao adotar o ticker como termo de busca, por entender que essa escolha fornece uma proxy mais precisa da atenção especificamente do investidor.

2.5 Proxies alternativas de atenção: redes sociais e plataformas de vídeo

O volume de buscas no Google é apenas uma das proxies possíveis para medir a atenção do investidor. A disseminação de plataformas digitais abriu espaço para outras alternativas, como as redes sociais e as plataformas de vídeo, que compartilham com o SVI o mesmo fundamento teórico — sinais do comportamento coletivo dos investidores carregam informação sobre o mercado —, mas diferem na fonte de dados.

Algumas dessas alternativas, inclusive, possibilitam levar em consideração o sentimento dos usuários e não somente a atenção. Bollen, Mao e Zeng (2011) utilizam o humor coletivo expresso no Twitter como proxy do estado emocional agregado da população. Por meio das ferramentas *OpinionFinder* e *Google-Profile of Mood States* (GPOMS), os autores extraem dimensões de humor a partir do conteúdo textual dos *tweets* e encontram que determinadas dimensões antecipam variações no índice Dow Jones Industrial Average.

Mais próximo da medida de atenção, Jang (2023) emprega a contagem de visualizações (*view count*) de vídeos no YouTube como proxy de atenção do investidor no mercado acionário sul-coreano. Ancorado na mesma literatura de atenção limitada (Hirshleifer; Teoh, 2003; Andrei; Hasler, 2015), o autor encontra que as visualizações de hoje antecipam a atenção, a compra por investidores de varejo e o retorno do dia seguinte, com efeito mais pronunciado entre empresas menores e no contexto posterior à pandemia de COVID-19. O resultado reforça, em outra plataforma e em um mercado emergente asiático, a generalidade do mecanismo que vincula atenção a retorno e volatilidade.

2.6 Posicionamento e contribuição deste trabalho

No que diz respeito à proxy de atenção, o presente estudo adota o *Search Volume Index* do Google Trends, na esteira de DA, ENGELBERG e GAO (2011) e de Perlin et al. (2017). Quanto ao termo de busca, alinha-se a DA, ENGELBERG e GAO (2011) ao utilizar o ticker de cada ação, diferentemente de Perlin et al. (2017), que recorrem a palavras genéricas de finanças. Em relação à variável-alvo, o foco recai sobre a volatilidade realizada, aproximando-se de Perlin et al. (2017), ao passo que DA, ENGELBERG e GAO (2011) concentram-se sobre o retorno.

Duas dimensões adicionais distinguem o trabalho. A primeira é o mercado analisado: enquanto a literatura-base concentra-se em mercados desenvolvidos, este estudo volta-se ao mercado acionário brasileiro, um mercado emergente ainda pouco explorado sob essa

ótica e marcado, na última década, por expressivo crescimento da base de investidores pessoa física na B3. A segunda é a abordagem econométrica: em vez de avaliar relações de causalidade por meio de modelos VAR e testes de Granger, como em Perlin et al. (2017), o presente trabalho avalia a acurácia preditiva fora da amostra, comparando um passeio aleatório, um modelo eGARCH(1,1) e um modelo eGARCH-X(1,1) que incorpora a atenção anormal de buscas (ASVI), por meio da função de perda QLIKE combinada ao procedimento Model Confidence Set. Trata-se de um teste sobre o valor preditivo incremental da atenção, pois não basta que a variável seja estatisticamente significativa, é preciso que sua inclusão melhore, de fato, a previsão da volatilidade.

Em síntese, há evidência consolidada de que a atenção captada pelo volume de buscas prevê a volatilidade em índices de mercados desenvolvidos (Perlin et al., 2017) e de que o SVI por ticker capta a atenção do investidor de varejo (DA; ENGELBERG; GAO, 2011). Permanece, contudo, em aberto a questão de saber se a atenção por ticker melhora a previsão da volatilidade, em um exercício preditivo fora da amostra, aplicado a um painel amplo de ações individuais de um mercado emergente como o brasileiro, lacuna que este trabalho se propõe a preencher.

3 DADOS E MÉTODO

Esta seção detalha os dados utilizados no estudo, fornecendo uma apresentação criteriosa de suas fontes. Em seguida, é delineada a metodologia adotada de maneira clara e rigorosa, em que são descritos os procedimentos estatísticos e as abordagens analíticas.

3.1 Dados

A base de dados do estudo é composta por duas fontes distintas: dados financeiros de preços e volume negociado de ações listadas na B3, obtidos via *Yahoo Finance*, e dados de atenção do investidor representados pelo *Search Volume Index* (SVI) do Google Trends. As duas séries são combinadas em uma única base, organizada no formato painel, com uma observação por par (semana, ativo). A base final compreende 260 semanas, da semana encerrada em 02 de abril de 2021 à semana encerrada em 20 de março de 2026, o que corresponde a aproximadamente cinco anos de dados históricos.

3.1.1 Coleta do Search Volume Index via Google Trends

O Google Trends é uma ferramenta pública que captura o interesse de busca no Google ao longo do tempo. Em vez de fornecer volumes absolutos de pesquisa, ele normaliza os dados em uma escala de 0 a 100: cada ponto é dividido pelo total de buscas da região e período selecionados, e o resultado é reescalonado proporcionalmente ao interesse relativo daquele termo em relação a todos os outros termos buscados. Isso significa que o valor 100 representa o pico de interesse relativo no período analisado, e não necessariamente

o maior volume absoluto de buscas em toda a existência do Google (Mavragani; Ochoa, 2019). A ferramenta ainda filtra buscas repetidas em curto intervalo, termos com volume muito baixo (reportados como 0) e caracteres especiais (Google, 2026), e os dados são agregados em granularidade semanal para períodos maiores que um ano e menores que cinco anos (Mavragani; Ochoa, 2019).

Apesar de sua utilidade para capturar a atenção do investidor, o Google Trends apresenta limitações relevantes. A normalização relativa impede comparações de magnitude entre períodos ou regiões diferentes: um SVI de 50 em janeiro não é diretamente comparável a um SVI de 50 em julho se o universo de buscas mudou. Em segundo lugar, atividades anômalas como buscas automatizadas podem permanecer nos dados como medida de segurança intencional do Google (Google, 2026). Por fim, a ferramenta não é uma pesquisa científica e não distingue a motivação por trás da busca (Google, 2026), de modo que um pico no SVI pode refletir tanto interesse de compra quanto preocupação ou curiosidade passageira do investidor.

Para melhor representar o interesse do investidor, os termos escolhidos foram os tickers das empresas listadas na bolsa de valores brasileira B3. O ticker foi escolhido para que não houvesse contaminação de pesquisas feitas por consumidores ou por interessados em notícias de uma determinada empresa. Por exemplo, quando é previsto um aumento no preço do petróleo, o termo "Petrobras" é mais pesquisado, porém não necessariamente por pessoas interessadas em investir na empresa. Isso vale também para consumidores pesquisando pelo site da Magazine Luiza ou Americanas para compras *online*. Ao escolher pelo ticker de cada empresa (Petrobras - PETR4, Magazine Luiza - MGLU3 e Americanas - AMER3), o ruído do aumento ou da queda por motivos não relacionados a investimentos diminui. Afinal, os resultados de pesquisa que aparecem ao buscar pelo ticker são, majoritariamente, de canais de investimentos, cotações, relações com investidor ou conteúdo educativo de investimento.

O SVI é coletado por meio da biblioteca *pytrends*, que oferece uma interface em *Python* para a API não oficial do Google Trends. O termo de busca utilizado é o próprio código do ticker da ação na B3, com a região delimitada ao Brasil (`geo = BR`) e as buscas são feitas individualmente. O SVI retornado é uma série semanal normalizada no intervalo de 0 a 100, em que o valor representa o pico ou o vale de busca relativo ao período solicitado. Logo, não é possível comparar o SVI entre os tickers, visto que 100 para PETR4 representa uma quantidade de pesquisas diferente de 100 para MGLU3.

3.1.2 Universo de ativos e critérios de elegibilidade

Os dados de preços diários são obtidos por meio da biblioteca *yfinance*. O universo inicial de ativos é composto por 387 ações listadas na B3. Para que um ativo seja incluído na amostra, dois critérios de elegibilidade são aplicados durante a etapa de coleta de dados financeiros.

O primeiro critério é de suficiência histórica: o ativo deve apresentar pelo menos 85% dos pregões esperados ao longo do horizonte amostral, calculados com base em aproximadamente 252 dias de negociação por ano. Ativos com histórico inferior a esse limite são excluídos, o que afasta papéis com listagem recente ou com longos períodos de suspensão de negociação.

O segundo critério é de liquidez mínima: o volume financeiro médio diário do ativo, calculado como o produto entre o número de ações negociadas e o preço de fechamento, deve ser superior a um piso mínimo de mil reais diários. Esse filtro tem o objetivo de excluir ativos com baixíssima atividade de negociação, para os quais a formação de preços pode ser irregular e as séries de volatilidade menos informativas.

Além dos dois critérios acima, um filtro adicional é aplicado após a agregação semanal: somente os ativos que apresentam o número completo de observações semanais são mantidos. Esse critério garante a homogeneidade do painel e elimina ativos com lacunas nas séries de preços. São coletadas as séries de preço de fechamento ajustado e de volume negociado para todos os ativos elegíveis.

O alinhamento entre as duas fontes de dados requer atenção especial ao calendário de referência de cada série. Os dados do Google Trends são organizados por semanas que iniciam no domingo e se encerram no sábado, ao passo que os dados financeiros são agregados por semanas iniciadas na segunda-feira e encerradas na sexta-feira. Para garantir a compatibilidade entre as duas bases, as datas da série de SVI são deslocadas de domingo para a sexta-feira imediatamente posterior, alinhando o identificador temporal das duas fontes, de modo a comparar o fechamento do preço e do SVI de cada ação. A base resultante é organizada no formato painel, com número fixo de observações por ativo e uma linha por par (semana, ticker).

3.2 Metodologia

O estudo adota um exercício de previsão fora da amostra para avaliar se a inclusão do *Search Volume Index* do Google Trends melhora a capacidade preditiva sobre a volatilidade das ações. A comparação é feita entre três modelos competidores, estimados com janela móvel de tamanho fixo e avaliados por uma função de perda robusta para variâncias combinada com o procedimento *Model Confidence Set* (MCS).

3.2.1 Volatilidade realizada

A volatilidade não é diretamente observável, de modo que a avaliação dos modelos requer uma *proxy* mensurável que represente a variação real do processo de preço (Andersen; Bollerslev, 1998). A medida adotada é a volatilidade realizada (RV), fundamentada na teoria de variação quadrática (Andersen; Bollerslev, 1998). Embora a teoria seja frequentemente aplicada a dados de alta frequência intradiária para estimar a volatilidade

diária, Andersen, Bollerslev et al. (2003) demonstram que o arcabouço é escalável para frequências mais baixas, como a agregação semanal a partir de retornos diários.

Assim, a volatilidade realizada é construída a partir dos retornos logarítmicos diários $r_{i,t} = \ln(P_{i,t}/P_{i,t-1})$ dentro de cada semana t:

$$RV_t = \sum_{i=1}^{n_t} r_{i,t}^2 \tag{1}$$

onde n_t é o número de pregões da semana t. A soma dos quadrados dos retornos diários fornece um estimador não enviesado e consistente da variância integrada do processo de preço (Andersen; Bollerslev et al., 2003).

A média dos retornos não é subtraída no cálculo da RV_t por duas razões teóricas centrais. Primeira, em horizontes temporais curtos (como o diário em relação ao semanal), a variação induzida pela volatilidade é de uma ordem de magnitude superior (ordem h) à variação causada pela média condicional (ordem h^2) (Andersen; Bollerslev et al., 2003). Assim, o impacto da variância é maior e mais notável que o impacto da média, ou seja, a variância é dominante. Segunda, como os retornos diários apresentam médias empiricamente próximas de zero, a sua manutenção simplifica a estimação e evita a introdução de ruído ou a perda desnecessária de graus de liberdade em amostras de tamanho limitado, o que poderia degradar a qualidade da *proxy* (Andersen; Bollerslev et al., 2003). Em outras palavras, em vez de utilizar apenas o preço de fechamento de cada semana, é melhor calcular a volatilidade utilizando os dados diários, pois estes retêm mais informação e, consequentemente, são estatisticamente mais eficientes.

3.2.2 A variável de atenção: construção e estacionariedade

Só é possível realizar previsões se o futuro for igual ao passado. Caso contrário, "essas relações históricas podem não ser guias confiáveis para o futuro" (Stock; Watson, 2004). Esse é o conceito central de estacionariedade. O futuro é igual ao passado. Por isso, modelos de séries temporais requerem que a série seja estacionária, condição necessária para que as estimativas dos parâmetros do modelo sejam válidas e as inferências estatísticas confiáveis. O uso de séries não estacionárias como regressores pode gerar relações espúrias e comprometer a validade das inferências.

Uma série temporal é considerada fracamente estacionária se três condições são atendidas: a média μ é constante ao longo do tempo; a variância σ^2 é finita e constante; e a covariância γ_s entre Y_t e Y_{t-s} é finita e depende apenas da defasagem s, e não do tempo absoluto t (Stock; Watson, 2004). As fontes mais importantes de não estacionariedade, conforme discutido por Stock e Watson (2004), são as tendências e as quebras estruturais. O SVI no nível bruto tende a exibir tendências e mudanças de patamar típicas de séries de popularidade de buscas, o que viola a hipótese de estacionariedade e, ao mesmo tempo, captura sobretudo o nível de interesse, e não a atenção anormal de interesse.

Por essas duas razões, a variável de atenção utilizada nos modelos é a atenção anormal de buscas, operacionalizada pelo desvio logarítmico do SVI em relação à sua mediana das oito semanas anteriores, assim como DA, ENGELBERG e GAO (2011):

$$\Delta_8 \text{SVI}_t = \ln(1 + \text{SVI}_t) - \ln\left(1 + \underset{k \in \{t-8, \ldots, t-1\}}{\text{med}} \text{SVI}_k\right) \tag{2}$$

A variável mede o quanto o interesse de busca da semana corrente está acima ou abaixo do seu nível "normal" recente, resumido pela mediana das oito semanas precedentes. A mediana, mais robusta a valores extremos do que a média, filtra o movimento lento de nível e isola o componente de atenção anormal. O uso da transformação $\ln(1 + \cdot)$ (*log1p*) evita a indefinição do logaritmo quando o SVI é igual a zero, situação frequente em ativos de menor liquidez, e estabiliza a variância da série. Como a medida depende da mediana das oito semanas anteriores, as oito primeiras semanas de cada série não possuem valor definido e são descartadas na construção da base final.

A estacionariedade de Δ_8SVI é verificada pelo teste de Dickey-Fuller aumentado (ADF), estimado por mínimos quadrados ordinários (MQO) com seleção automática de defasagens pelo critério de Akaike, a partir da Equação (3):

$$\Delta Y_t = \beta_0 + \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \gamma_2 \Delta Y_{t-2} + \cdots + \gamma_p \Delta Y_{t-p} + u_t \tag{3}$$

$$\text{Teste ADF} = \begin{cases} H_0: & \delta = 0 \\ H_1: & \delta < 0 \end{cases}$$

Sob a hipótese nula H_0, a série Y_t possui raiz unitária (tendência estocástica); sob a hipótese alternativa H_1, a série é estacionária. A estatística ADF é a estatística t unicaudal de MQO associada ao coeficiente δ. Entretanto, essa estatística não segue uma distribuição normal, mesmo em amostras grandes, e seus valores críticos são obtidos a partir da tabela específica do teste calculada por Dickey e Fuller (1979).

Nos testes realizados, o nível de significância adotado foi de 5%. O SVI no nível bruto apresentou estacionariedade em apenas 207 dos 308 tickers testados, confirmando a presença frequente de raiz unitária na série de nível. Já a atenção anormal Δ_8SVI apresentou estacionariedade em 302 dos 308 tickers. Os seis ativos cuja série Δ_8SVI é classificada como não estacionária pelo teste ADF são retirados da base final, de modo que a variável exógena utilizada nos modelos de previsão satisfaça os requisitos de estacionariedade.

3.2.3 Exercício de previsão fora da amostra com janela móvel

Para simular um ambiente realista de previsão, os modelos são estimados e avaliados em um exercício fora da amostra com janela móvel de tamanho fixo de 156 semanas (equivalente a três anos de pregões). O procedimento é o seguinte: para cada semana t do

período de teste, os parâmetros de cada modelo são reestimados utilizando as 156 semanas imediatamente anteriores; em seguida, gera-se a previsão de variância $\hat{\sigma}_t^2$ para a semana t; a janela avança uma semana e o processo se repete até o final da amostra.

Esse procedimento garante que nenhuma informação futura seja incorporada à estimação ou à previsão, prevenindo o viés de antecipação (*look-ahead bias*). O alinhamento temporal entre as séries é o seguinte: o valor $\Delta_8 \mathrm{SVI}_{t-1}$ utilizado na previsão para a semana t corresponde à atenção anormal encerrada na semana $t-1$, a última observação disponível antes do início dos pregões da semana t. Como a janela de treino exige 156 semanas e o regressor exógeno demanda uma defasagem adicional, a primeira previsão recai sobre a 158ª semana da série filtrada. O exercício gera, assim, 103 previsões fora da amostra por ativo, cobrindo o período da semana de 05 de abril de 2024 a de 20 de março de 2026. Os modelos M0, M1 e M2 são estimados nas mesmas semanas-calendário para cada ticker, em bloco contíguo, garantindo comparabilidade entre modelos e entre ativos.

3.2.4 Modelo 0: passeio aleatório

O passeio aleatório (*random walk*) constitui o modelo mais simples capaz de representar uma variável com tendência estocástica. Sua formulação parte da ideia de que o valor de uma série no período t corresponde ao seu valor no período anterior acrescido de um choque imprevisível, de modo que o futuro da série depende inteiramente do seu estado presente e de uma perturbação aleatória. Formalmente, o modelo é expresso como:

$$Y_t = Y_{t-1} + u_t \tag{4}$$

onde u_t é um termo de erro com distribuição independente e identicamente distribuída (i.i.d.), de média zero e variância constante σ^2. O coeficiente que multiplica Y_{t-1} é igual a 1, o que caracteriza a presença de raiz unitária no processo. De acordo com a hipótese de mercados eficientes, os preços das ações adotam o passeio aleatório, sendo impossível prever o preço de amanhã (Gujarati; Porter, 2011).

O Modelo 0 (M0) é utilizado como referência ingênua. Ele assume que a melhor previsão para a variância da próxima semana é a volatilidade realizada da semana corrente:

$$\hat{\sigma}_t^2 = RV_{t-1} \tag{5}$$

Esse modelo não possui parâmetros a estimar e serve como piso de desempenho: qualquer modelo que não supere o M0 não apresenta valor preditivo.

3.2.5 Modelo 1: eGARCH

O Modelo 1 corresponde à especificação eGARCH(1,1) (*exponential GARCH*), introduzida por Nelson (1991) como aprimoramento do GARCH(1,1) de Bollerslev (1986). Os modelos da família GARCH capturam dois fatos estilizados centrais das séries de

retornos financeiros: a heterocedasticidade condicional e o agrupamento de volatilidade, pelo qual períodos de alta variância tendem a ser seguidos por períodos de alta variância, e analogamente para períodos calmos. O eGARCH adiciona dois aperfeiçoamentos importantes. Primeiro, ao modelar o logaritmo da variância condicional, dispensa as restrições de não negatividade sobre os parâmetros, pois a exponenciação garante variância positiva por construção. Este é um ponto importante, visto que a variável $\Delta_8\mathrm{SVI}_{t-1}$ adota valores negativos. Segundo, captura o efeito alavancagem, isto é, a assimetria empírica pela qual choques negativos nos retornos elevam mais a volatilidade futura do que choques positivos de mesma magnitude.

A equação da média é especificada como:

$$r_t = \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t, \quad z_t \sim N(0,1) \tag{6}$$

onde r_t denota o retorno semanal logarítmico, μ é a média condicional constante, ε_t é a inovação condicionalmente heterocedástica, σ_t^2 é a variância condicional no instante t e z_t é um processo i.i.d. com distribuição normal padronizada. A variância condicional evolui segundo a equação de log-variância:

$$\ln \sigma_t^2 = \omega + \alpha\, z_{t-1} + \xi\,(|z_{t-1}| - E|z_{t-1}|) + \beta \ln \sigma_{t-1}^2 \tag{7}$$

O termo $\alpha\, z_{t-1}$ captura o efeito de *sinal* (alavancagem): um α negativo indica que choques negativos elevam mais a volatilidade do que choques positivos. O termo $\xi\,(|z_{t-1}| - E|z_{t-1}|)$ captura o efeito de *magnitude*: a reação da variância ao tamanho do choque, independentemente do sinal. O coeficiente β mede a persistência da volatilidade, isto é, a velocidade com que a memória de choques passados se dissipa; a estacionariedade do processo de log-variância exige $|\beta| < 1$, sem necessidade de restrições adicionais de positividade (Nelson, 1991).

A estimação do vetor de parâmetros é realizada por máxima verossimilhança, por meio do pacote *rugarch* (Ghalanos, 2024) da linguagem *R*, com otimização numérica híbrida. Para fins de previsão fora da amostra em horizonte de um passo à frente, a previsão da variância condicional para $t+1$ é obtida da equação recursiva avaliada nos parâmetros estimados na janela corrente, e os parâmetros são reestimados a cada passo da janela móvel descrita na Seção 3.2.3, de modo que a previsão para a semana $t+1$ utiliza apenas informação disponível até a semana t. Na implementação, os retornos semanais logarítmicos são previamente reescalados pelo fator 100, prática usual destinada a evitar problemas de convergência numérica frequentes em séries financeiras com magnitudes próximas a zero. Após a estimação e a geração das previsões, as variâncias condicionais são reescalonadas pelo fator $1/10\,000$, a fim de retornar à escala original e permitir a comparação direta com a volatilidade realizada construída na Seção 3.2.1.

3.2.6 Modelo 2: eGARCH-X

O eGARCH(1,1) descreve a dinâmica da volatilidade usando apenas informação contida na própria série de retornos, isto é, choques passados e variâncias passadas. Essa estrutura captura bem o agrupamento de volatilidade, mas é, por construção, fechada em si mesma. Nada que ocorra fora da série pode entrar no modelo. A extensão eGARCH-X relaxa essa restrição ao permitir que uma variável externa, observável em $t-1$, ajude a explicar a variância condicional em t. O Modelo 2 incorpora a atenção anormal de buscas diretamente na equação de log-variância como regressor exógeno:

$$\ln \sigma_t^2 = \omega + \alpha\, z_{t-1} + \xi\,(|z_{t-1}| - E|z_{t-1}|) + \beta \ln \sigma_{t-1}^2 + \gamma\, \Delta_8 \mathrm{SVI}_{t-1} \tag{8}$$

Os quatro primeiros termos preservam toda a estrutura do eGARCH(1,1) padrão. O termo adicional $\gamma\, \Delta_8 \mathrm{SVI}_{t-1}$ é o objeto central de interesse: ele mede o efeito marginal da atenção anormal do investidor sobre a log-variância condicional, condicional à dinâmica que o eGARCH já captura. Por entrar na equação do logaritmo da variância, γ tem interpretação de semielasticidade: indica a variação proporcional da variância condicional associada a uma unidade de atenção anormal.

Sob a hipótese nula $\gamma = 0$, a atenção do investidor é redundante uma vez controlados os efeitos próprios da série de retornos, e o eGARCH-X equivale ao eGARCH padrão. Um γ positivo e estatisticamente significativo indicaria que picos de atenção anormal antecedem picos de volatilidade, consistente com a tese de que o aumento de atenção precede ou acompanha movimentos mais turbulentos no preço. O pacote *rugarch* permite incluir o regressor diretamente na equação da variância, de modo que γ mede o efeito da atenção sobre a volatilidade, e não sobre o retorno esperado. A comparação entre M1 e M2 é feita exclusivamente em termos de desempenho preditivo agregado, via QLIKE e *Model Confidence Set*.

3.2.7 Identificação de γ e tratamento das séries de SVI

A estimação do eGARCH-X por máxima verossimilhança depende da inversão da matriz Hessiana da função de log-verossimilhança, que reúne as segundas derivadas parciais em relação a cada parâmetro do modelo. Essa inversão guia as iterações do algoritmo de otimização e, após a convergência, fornece a base para o cálculo da matriz de covariância assintótica dos estimadores. A existência da inversa pressupõe que a Hessiana seja não singular; uma matriz singular não admite inversa, inviabilizando tanto a atualização dos parâmetros quanto o cálculo dos erros padrão.

O parâmetro γ só é identificável se o regressor Δ_8SVI apresentar variação suficiente dentro da janela de estimação. Quando a série é constante - por exemplo, quando o SVI permanece em zero durante toda a janela -, a atenção anormal não varia, a superfície da log-verossimilhança torna-se plana na direção de γ e a Hessiana torna-se singular, fazendo a estimação falhar. Em situações intermediárias, nas quais a série apresenta excesso de

zeros sem ser inteiramente constante, o modelo converge, porém os estimadores ficam sujeitos a distorções.

Essas distorções decorrem de duas fontes. A primeira é um erro de medição na variável explicativa: o Google Trends trunca para zero qualquer volume de busca abaixo de um limiar interno, de modo que os zeros observados podem refletir tanto ausência real de atenção quanto atenção positiva, porém abaixo do limite de detecção. A segunda é o *overfitting* sobre os poucos pontos não nulos: uma janela de 156 observações com poucos valores distintos de zero fornece informação escassa para identificar γ, elevando a variância amostral do estimador e tornando-o sensível a flutuações idiossincráticas da janela corrente, o que degrada a qualidade preditiva fora da amostra.

Diante disso, dois tratamentos são adotados. Primeiro, como descrito na Seção 3.2.2, os ativos cuja série Δ_8SVI é não estacionária são excluídos. Segundo, na consolidação dos resultados, descartam-se os tickers cujas janelas do M2 produzem previsões numericamente patológicas (não finitas ou abaixo de um piso mínimo de variância), que sinalizam falha do ajuste. Como a sensibilidade dos resultados à presença de séries com muitos zeros poderia ser objeto de dúvida, a robustez da análise é avaliada na Seção 4, comparando esse critério principal com a remoção adicional dos tickers que incorrem em erro de estimação (Hessiana singular) e com diferentes limiares máximos para a proporção de zeros do SVI.

3.2.8 Função de perda QLIKE

O desempenho preditivo dos modelos é medido pela função de perda quase-verossimilhança (QLIKE), definida para cada semana t como:

$$L_t\left(\hat{\sigma}_t^2, RV_t\right) = \frac{RV_t}{\hat{\sigma}_t^2} - \ln\left(\frac{RV_t}{\hat{\sigma}_t^2}\right) - 1 \tag{9}$$

A QLIKE é especialmente adequada para comparar previsões de variância porque apresenta a propriedade de robustez ao uso de *proxies* ruidosas: quando a medida observada de volatilidade (RV_t) não é a variância verdadeira, mas sim uma estimativa consistente desta, a ordenação dos modelos pela QLIKE média converge para a ordenação obtida com a variância latente verdadeira. Essa propriedade, demonstrada por Patton (2011), distingue a QLIKE do erro quadrático médio, que não possui a mesma robustez. Modelos com menor QLIKE média apresentam previsões mais próximas da volatilidade realizada observada. A função é assimétrica, penalizando mais fortemente as previsões que subestimam a volatilidade.

3.2.9 Model Confidence Set

A comparação de modelos de previsão enfrenta uma dificuldade fundamental: em qualquer amostra finita, raramente os dados são suficientemente informativos para apontar de forma inequívoca um único modelo como superior a todos os concorrentes. Diante disso,

Hansen, Lunde e Nason (2011) propõem o *Model Confidence Set* (MCS), um procedimento que, em vez de eleger um vencedor isolado, constrói um conjunto de modelos que contém o melhor com uma probabilidade pré-especificada, de forma análoga a um intervalo de confiança para um parâmetro populacional. Dados informativos produzem um MCS reduzido, possivelmente unitário, ao passo que dados pouco informativos produzem um MCS amplo, refletindo a incapacidade da amostra de discriminar entre as alternativas. Formalmente, parte-se de um conjunto inicial de modelos candidatos $\mathcal{M}_0$ e define-se o conjunto de modelos superiores como

$$\mathcal{M}^* \equiv \{i \in \mathcal{M}_0 : \mu_{ij} \leq 0 \text{ para todo } j \in \mathcal{M}_0\}, \tag{10}$$

onde $\mu_{ij} \equiv E(L_{i,t} - L_{j,t})$ representa a diferença esperada de perda entre os modelos i e j. O objetivo do procedimento é determinar um conjunto $\mathcal{M}^*_{1-\alpha}$ tal que $P(\mathcal{M}^* \subset \mathcal{M}^*_{1-\alpha}) \geq 1-\alpha$ assintoticamente.

O algoritmo é iterativo: a cada passo, aplica-se um teste de equivalência à hipótese nula $H_{0,\mathcal{M}} : \mu_{ij} = 0$ para todo $i, j \in \mathcal{M}$; se rejeitada, uma regra de eliminação remove o pior modelo e o procedimento se repete até a primeira aceitação. Como a distribuição assintótica da estatística de teste depende da matriz de correlação entre as perdas relativas, parâmetro de incômodo de difícil estimação direta, Hansen, Lunde e Nason (2011) propõem uma implementação por *bootstrap* em blocos (*moving-block bootstrap*), que preserva a estrutura de dependência temporal das séries de perda. Para cada modelo, calcula-se um *MCS p-value*, de modo que o modelo i pertence a $\mathcal{M}^*_{1-\alpha}$ se e somente se seu *p-value* for maior ou igual a α.

Neste trabalho, o MCS é aplicado para comparar a acurácia preditiva fora da amostra dos três modelos competidores (M0, M1 e M2). A função de perda utilizada é a QLIKE, computada semanalmente, conforme detalhado na Seção 3.2.8. O procedimento é implementado por meio do módulo *MCS* da biblioteca *arch* em *Python*, com nível de significância $\alpha = 0{,}10$ (resultando em $\mathcal{M}^*_{90\%}$), tamanho de bloco igual a 5 semanas e $B = 1\,000$ reamostragens *bootstrap*. O tamanho de bloco foi escolhido em consonância com a frequência semanal dos dados e segue a recomendação geral do artigo original de calibrar a janela do *bootstrap* à estrutura de dependência da série.

4 RESULTADOS

Esta seção apresenta os resultados do exercício de previsão fora da amostra. Inicialmente, descreve-se a composição final do painel de ativos; em seguida, comparam-se os três modelos quanto à acurácia preditiva, por meio da QLIKE e do *Model Confidence Set*; analisa-se o papel da variável de atenção; e, por fim, avalia-se a robustez dos resultados a diferentes critérios de filtragem das séries de SVI.

4.1 Composição do painel

Do conjunto de 308 tickers com séries de SVI coletadas e submetidas ao teste ADF, 302 apresentaram a variável de atenção Δ_8SVI estacionária e foram efetivamente estimados nos três modelos. Desses, 72 incorreram em erro de estimação no Modelo 2 (Hessiana singular ou falha de convergência) em ao menos uma janela, padrão associado a séries de SVI com proporção elevada de zeros, conforme antecipado na Seção 3.2.7. A análise principal adota o critério mais conservador de não impor um limiar máximo de zeros, descartando apenas os ativos com previsões numericamente patológicas; o painel final reúne, assim, 260 tickers e 26.780 observações fora da amostra.

4.2 Acurácia preditiva: QLIKE e Model Confidence Set

A Tabela 1 resume o desempenho preditivo dos três modelos. O passeio aleatório (M0) apresenta a maior perda, sendo claramente dominado pelos modelos da família GARCH. Entre estes, o eGARCH puro (M1) registra a menor QLIKE, tanto na média agrupada sobre as 25.405 observações quanto na mediana entre tickers, e é o modelo com melhor desempenho individual no maior número de ativos (182 dos 260).

Tabela 1 – Desempenho preditivo fora da amostra dos três modelos.

Modelo	**QLIKE médio**	**QLIKE mediano**	**Melhor em (tickers)**	**MCS (p-valor)**
M0 - Passeio aleatório	39,151	1,655	9	0,039 (fora)
M1 - eGARCH(1,1)	3,205	0,691	182	1,000 (dentro)
M2 - eGARCH-X(1,1)	8,819	0,768	69	0,092 (fora)

Fonte: elaboração própria (2026). **Nota:** QLIKE médio agrupado sobre as 25.405 observações fora da amostra; QLIKE mediano calculado entre os 260 tickers. "Melhor em" indica o número de tickers em que cada modelo registra a menor QLIKE média. p-valor do *Model Confidence Set* a $\alpha = 0{,}10$.

No resultado do *Model Confidence Set* ao nível de significância de 10%, o conjunto de confiança $\mathcal{M}^*_{90\%}$ contém apenas o Modelo 1 (p-valor = 1,000). O eGARCH-X com atenção anormal (M2) é eliminado com p-valor de 0,092 e o passeio aleatório (M0), com p-valor de 0,039. Em outras palavras, ainda que ambos os modelos GARCH superem amplamente o *benchmark* ingênuo, a inclusão da variável de atenção não apenas deixa de melhorar a previsão como a deteriora no agregado, levando à exclusão do M2 do conjunto dos modelos estatisticamente indistinguíveis do melhor.

A QLIKE média agrupada é fortemente influenciada por um pequeno número de tickers com perdas extremas, o que explica os valores elevados das médias (sobretudo do M0). A mediana entre tickers, mais robusta a esses casos, confirma de forma consistente a ordenação M1 (0,691) < M2 (0,768) < M0 (1,655), reforçando a superioridade do eGARCH puro.

4.3 Diagnóstico dos resíduos

A adequação das especificações GARCH é avaliada pelos resíduos padronizados $z_t = \varepsilon_t/\sigma_t$ do último ajuste de cada janela. Aplicam-se o teste de Ljung-Box sobre z_t (autocorrelação remanescente na média) e sobre z_t^2 (efeitos ARCH não capturados), além do teste de Jarque-Bera (normalidade). Os resultados indicam que as especificações eGARCH e eGARCH-X capturam adequadamente a dinâmica de heterocedasticidade da maior parte dos ativos, ainda que a hipótese de normalidade das inovações seja rejeitada em parcela relevante das séries, padrão usual em retornos de ações e que sugere, como extensão, o uso de distribuições com caudas mais pesadas. O resultado de todos os tickers estão no Apêndice A.

4.4 Robustez à filtragem das séries de SVI

Como a proporção de zeros nas séries de SVI afeta a identificação de γ (Seção 3.2.7), avalia-se a robustez dos resultados a diferentes critérios de filtragem. A Tabela 2 reporta a QLIKE média e o conjunto de confiança do MCS sob quatro cenários: sem filtro de zeros (análise principal), com remoção adicional dos tickers que incorrem em erro de estimação no M2, e com a exclusão de ativos cuja série de SVI tem 70% ou mais e 50% ou mais de semanas nulas.

Tabela 2 – Robustez do *Model Confidence Set* a critérios de filtragem das séries de SVI.

Critério (tickers)	**QLIKE M0**	**QLIKE M1**	**QLIKE M2**	**MCS** ($\alpha = 0{,}10$)
Sem filtro de zeros (260)	39,151	3,205	8,819	{M1}
Sem erro de estimação (230)	37,439	0,894	1,011	{M1}
$\geq 70\%$ de zeros (173)	1,738	0,758	0,849	{M1}
$\geq 50\%$ de zeros (156)	1,714	0,760	0,829	{M1}

Fonte: elaboração própria (2026). **Nota:** QLIKE média agrupada por cenário. Em todos os critérios, o conjunto de confiança do MCS a 10% contém exclusivamente o Modelo 1 (eGARCH puro).

Em todos os cenários, o conjunto de confiança do MCS contém exclusivamente o Modelo 1 (eGARCH puro). Um comportamento empírico notável revelado na Tabela 2 é a redução drástica das perdas médias não apenas do Modelo 2, mas também dos Modelos 0 (passeio aleatório) e 1 (eGARCH puro), que sequer utilizam a variável de atenção em suas formulações. Ao transitar do cenário sem filtro para o filtro de zeros $\geq 70\%$, a perda QLIKE média do passeio aleatório desaba de $39,151$ para $1,738$, e a do eGARCH cai de $3,205$ para $0,758$.

Esse fenômeno revela que a proporção de zeros nas séries de SVI funciona, de maneira indireta, como proxy indireta para a liquidez e para a estabilidade de mercado dos

ativos. Ações que registram buscas nulas em grande parte das semanas são tipicamente de baixíssima liquidez. Esses ativos apresentam dinâmicas de preços extremamente erráticas, marcadas por longos períodos de inatividade comercial interrompidos por saltos abruptos de preço. Modelar e prever a volatilidade sob tais condições é intrinsecamente desafiador para qualquer especificação econométrica, o que se traduz em erros de previsão extremos e eleva a perda média agrupada. Assim, ao filtrar séries com excesso de zeros no Google Trends, o pesquisador está, implicitamente, saneando a amostra de ativos com dinâmica de volatilidade patológica. Contudo, mesmo restringindo o exercício preditivo ao conjunto de ações mais líquidas e com maior densidade de sinal de atenção, o eGARCH-X não recupera a paridade com o eGARCH puro, o que atesta a robustez de sua redundância preditiva fora da amostra.

4.5 Sobre os ativos com SVI significativo

A divergência entre a significância estatística do coeficiente γ e a sua utilidade preditiva fora da amostra, observada na Seção 4.2, merece exame em profundidade. Por que o eGARCH-X (M2) não supera o eGARCH puro (M1) apesar de o coeficiente da atenção ser estatisticamente significativo em parcela expressiva dos ativos?

A Tabela 3 reporta a volatilidade realizada (RV) dos ativos agrupados segundo o modelo vencedor por QLIKE. Os 69 ativos em que o M2 vence apresentam RV mediano de $3{,}01 \times 10^{-3}$ e volatilidade anualizada mediana de 39,6%, valores próximos dos 182 ativos vencidos pelo M1 (RV mediano $3{,}25 \times 10^{-3}$; volatilidade anualizada 41,1%). Concentração de RV extrema observa-se apenas nos 9 ativos vencidos pelo passeio aleatório (M0), com RV mediano $7{,}79 \times 10^{-3}$ e volatilidade anualizada de 63,6%, perfil característico de microcaps em colapso. Em resumo, a vitória do M2 sobre o M1, em ativos individuais, não se associa a um perfil distinto de volatilidade.

Tabela 3 – Volatilidade realizada (RV) dos ativos segundo o modelo vencedor pelo QLIKE. Os ativos em que o M2 vence não se distinguem em volatilidade dos demais (Mann–Whitney $p = 0,72$); apenas os poucos casos vencidos pelo passeio aleatório (M0) concentram volatilidade extrema.

Modelo vencedor	N	RV médio ($\times 10^{-3}$)	RV mediano ($\times 10^{-3}$)	Vol. anual. mediana (%)
M0 (RW)	9	6,39	7,79	63,6
M1 (eGARCH)	182	4,41	3,25	41,1
M2 (eGARCH-X)	69	4,20	3,01	39,6

Fonte: elaboração própria (2026).

A Tabela 4 desloca o foco do vencedor de QLIKE para a robustez in-sample do coeficiente γ, estratificando 259 ativos em faixas de 25 pontos percentuais do indicador "percentual de janelas em que γ é significativo a 5%". A volatilidade anualizada mediana

é praticamente indistinguível nas três primeiras faixas — 41,0%, 39,5% e 39,4%, respectivamente —, sem qualquer tendência de crescimento à medida que a significância de γ aumenta, e só então dá um salto abrupto para 52,2% na faixa superior, em que γ é significativo em pelo menos 75% das janelas. Em outras palavras, a volatilidade elevada não acompanha gradualmente a significância da atenção: trata-se de um fenômeno restrito ao quartil de significância mais alta — um efeito de cauda, e não uma relação contínua entre atenção e volatilidade.

Tabela 4 – Volatilidade realizada (RV) por faixa de significância do coeficiente de atenção (γ), em intervalos fixos de 25 pontos percentuais (percentual de janelas com γ significativo a 5%).

γ signif. (% das janelas)	N	RV médio ($\times 10^{-3}$)	RV mediano ($\times 10^{-3}$)	Vol. anual. mediana (%)
0–25	97	3,91	3,23	41,0
25–50	61	4,49	3,00	39,5
50–75	65	4,27	2,98	39,4
75–100	36	5,99	5,24	52,2

Fonte: elaboração própria (2026). **Nota:** um dos 260 ativos do painel é omitido por apresentar estimação singular do Modelo 2 em todas as janelas, o que deixa indefinido o indicador de significância de γ; assim, N totaliza 259.

Para compreender a dinâmica subjacente que explica a redundância preditiva do SVI, a Tabela 5 apresenta a correlação de Spearman mediana entre a atenção anormal (Δ_8SVI) e a volatilidade realizada (RV) em três horizontes temporais distintos. Observa-se que a associação contemporânea (SVI_t, RV_t) é a mais expressiva, atingindo $0,302$ para os ativos em que o coeficiente de atenção mostrou-se frequentemente significativo. Relevante notar que a correlação da atenção reagindo à volatilidade passada (SVI_t, RV_{t-1}) é de $0,121$, enquanto a correlação da atenção antecedendo a volatilidade futura (SVI_{t-1}, RV_t) praticamente desaparece ($0,030$). A correlação praticamente nula do SVI defasado com a volatilidade realizada futura é o resultado central deste ponto: a atenção de busca não traz informação preditiva nova sobre a volatilidade, pois toda a sua associação com o RV concentra-se no presente e no passado, e não no futuro. Assim, o fluxo de buscas no Google Trends constitui uma resposta tardia dos investidores de varejo a turbulências já ocorridas no mercado, e não um sinal antecedente; como o conteúdo informativo de SVI_{t-1} sobre RV_t é desprezível, ele nada acrescenta à memória autorregressiva que o modelo eGARCH já capta — o que explica a ausência de ganho preditivo incremental do eGARCH-X.

O conteúdo preditivo do SVI defasado sobre a volatilidade está, portanto, integralmente contido na própria memória do processo de volatilidade. Assim, a atenção captada pelo SVI constitui uma resposta ao processo de volatilidade, e não um sinal que o antecede; como o eGARCH puro já internaliza essa dinâmica por meio do termo de persistência $\beta \ln \sigma^2_{t-1}$ e do termo de magnitude $\xi\left(|z_{t-1}| - E|z_{t-1}|\right)$ da equação de log-variância (Nelson, 1991), o SVI defasado torna-se redundante e não agrega conteúdo preditivo incremental.

Tabela 5 – Correlação de Spearman (mediana intra-ativo) entre a atenção (Δ_8SVI) e a volatilidade realizada (RV). A atenção co-move com a volatilidade contemporânea e a segue, em vez de antecedê-la.

Correlação	Significantes	Demais
Contemporânea (SVI_t, RV_t)	0,302	0,198
Atenção segue (SVI_t, RV_{t-1})	0,121	0,080
Atenção antecede (SVI_{t-1}, RV_t)	0,030	0,036

Fonte: elaboração própria (2026). **Nota:** subconjunto "significantes" = 32 ativos com γ significativo a 5% em $\geq 70\%$ das janelas e $< 70\%$ de semanas com SVI igual a zero.

Em outras palavras, essa significância na amostra não se converte em ganho preditivo fora da amostra, porque a atenção é um reflexo do próprio processo de volatilidade que se pretende prever. O eGARCH puro já internaliza o mecanismo de "volatilidade gera volatilidade" por meio da persistência, e o SVI, sendo redundante com essa memória, não adiciona informação. Tal redundância oferece a interpretação econômica subjacente à exclusão do M2 do *Model Confidence Set* (Seção 4.2) e dialoga com a evidência de DA, ENGELBERG e GAO (2011) segundo a qual o SVI captura sobretudo a atenção do investidor de varejo, em geral reativa a movimentos já observados nos preços.

5 CONCLUSÃO

Este trabalho investigou se a atenção do investidor, captada pelo *Search Volume Index* do Google Trends para os tickers de ações listadas na B3, possui poder preditivo sobre a volatilidade realizada semanal desses ativos. Para responder a essa questão, conduziu-se um exercício de previsão fora da amostra, com janela móvel de 156 semanas sobre um painel de 260 ações, comparando três modelos competidores - um passeio aleatório (M0), um eGARCH(1,1) (M1) e um eGARCH-X(1,1) que incorpora a atenção anormal de buscas na equação da variância condicional (M2) - por meio da função de perda QLIKE e do procedimento *Model Confidence Set*.

A principal conclusão é que a atenção anormal de buscas não melhora a previsão da volatilidade no agregado. Embora ambos os modelos GARCH superem amplamente o passeio aleatório, o *Model Confidence Set* a 10% contém exclusivamente o eGARCH puro. O eGARCH-X que inclui o SVI é eliminado do conjunto, e esse resultado mostra-se robusto a todos os critérios de filtragem das séries de busca avaliados. A inclusão da variável de atenção, longe de aprimorar a previsão, tende a introduzir ruído que deteriora a acurácia preditiva fora da amostra, sobretudo em ativos com séries de SVI dominadas por zeros.

Ao mesmo tempo, os resultados não autorizam descartar por completo o conteúdo informacional da atenção. O coeficiente associado ao SVI é estatisticamente significativo em pelo menos metade das janelas para 101 dos 260 tickers, e o eGARCH-X (M2) vence

individualmente o eGARCH (M1) em 69 ativos. O achado mais relevante nesse ponto é de natureza metodológica: a significância estatística do coeficiente de atenção e a sua utilidade preditiva fora da amostra são fenômenos empiricamente independentes. Um regressor estatisticamente significativo na amostra de estimação não se traduz, necessariamente, em ganho de previsão - advertência pertinente à literatura que infere relevância preditiva da atenção a partir de testes de significância dentro da amostra.

Esses resultados contribuem para a literatura de finanças comportamentais e atenção do investidor em quatro pontos distintos. Primeiro, transpõem o teste para um mercado emergente ainda pouco explorado sob essa ótica, o brasileiro, marcado na última década por forte crescimento da base de investidores pessoa física na B3. Segundo, adotam uma medida de atenção alinhada à teoria - a atenção anormal de buscas por ticker, na esteira de DA, ENGELBERG e GAO (2011) -, em vez de termos genéricos de finanças. Terceiro, submetem a hipótese a um teste preditivo exigente, em que não basta a variável ser significativa, é preciso que sua inclusão melhore, de fato, a previsão da volatilidade. Quarto, evidenciam que a atenção do investidor no mercado brasileiro é predominantemente reativa, seguindo a volatilidade em vez de antecedê-la, o que explica a redundância do regressor no modelo preditivo.

O estudo apresenta limitações que apontam caminhos para pesquisas futuras. A medida de atenção herda as restrições do Google Trends, em especial o truncamento de baixos volumes de busca a zero, que introduz erro de medição e prejudica a identificação do coeficiente de atenção nos ativos menos buscados. Outra limitação decorre da granularidade mensal disponibilizada pelo Google Trends para períodos maiores que cinco anos. A especificação supõe inovações com distribuição normal, hipótese rejeitada em parcela das séries; o emprego de distribuições com caudas mais pesadas, como a t de Student ou suas versões assimétricas, constitui extensão natural. Além disso, o exercício restringe-se a previsões de um passo à frente e a uma única *proxy* de atenção. Investigações futuras podem explorar horizontes de previsão mais longos reindexando o SVI, medidas alternativas de atenção extraídas de redes sociais e plataformas de vídeo, a heterogeneidade do efeito entre setores e portes de empresa, e medidas de volatilidade realizada baseadas em dados intradiários de maior frequência.

1 APÊNDICE A - Resultados das estimações por ticker

Este apêndice reporta, para cada um dos 302 tickers estimados, os parâmetros médios ao longo das janelas móveis do exercício fora da amostra. As colunas $\hat{\gamma}$ Médio e p-valor Médio referem-se ao coeficiente da variável de atenção (Δ_8SVI) na equação de log-variância do Modelo 2 (eGARCH-X); a coluna % Signif. 5% indica a fração de janelas em que esse coeficiente é significativo a 5%; as colunas $\hat{\alpha} + \hat{\beta}$ reportam a soma dos coeficientes de sinal e de memória dos Modelos 1 e 2; e SVI Zeros (%) indica a proporção de semanas com SVI nulo para o ticker.

Tabela 6 – Resultados das estimações por ticker.

Ticker B3	**N. Prev.**	$\hat{\gamma}$ **Médio**	**p-valor Méd.**	**% Signif.**	$\hat{\alpha} + \hat{\beta}$ **M1**	$\hat{\alpha} + \hat{\beta}$ **M2**	**SVI Zeros**
AALR3	103	0,1012	0,2948	32,04	0,3101	0,4374	61,57
ABCB4	103	-0,2077	0,3834	16,50	0,8719	0,0880	0,00
ABEV3	103	0,1755	0,1898	40,78	0,4577	0,5495	0,00
AERI3	103	-0,4351	0,1532	49,51	0,8129	0,8752	1,87
AFLT3	103	0,3288	0,2873	0,00	0,4861	0,5087	99,63
AGRO3	103	0,2976	0,4139	13,59	0,9844	0,5715	0,00
AHEB3	103	-3,8111	0,0222	94,12	-0,0015	-0,4387	99,25
ALOS3	103	-0,1770	0,1694	37,86	0,6335	0,3143	52,61
ALPA3	103	-0,0355	0,2756	49,51	-0,0144	-0,2819	96,27
ALPA4	103	0,2055	0,0846	83,50	0,3819	0,3654	6,72
ALPK3	103	-0,1139	0,0228	91,26	0,3047	0,7018	63,43
ALUP11	103	0,0550	0,2738	49,51	0,7865	0,6585	0,37
ALUP3	103	0,1423	0,5760	24,27	0,3265	0,2473	69,03
ALUP4	103	0,6030	0,2132	62,14	0,8924	0,6760	29,85
AMAR3	103	0,0808	0,0845	84,47	0,7863	0,5698	4,10
AMBP3	103	-0,2542	0,1881	51,46	0,6021	0,6028	0,00
AMER3	103	0,0792	0,3488	28,16	0,7033	0,5971	8,96
ANIM3	103	-0,2617	0,3870	0,00	0,3242	0,3841	1,12
AXIA3	103	0,0304	0,1665	63,16	0,7725	0,7926	92,16
AXIA6	103	-0,3209	0,2002	22,58	0,7686	0,5334	92,54
AZEV3	103	0,0627	0,4381	11,65	0,9269	0,8144	32,09
AZEV4	103	0,1307	0,4723	12,62	0,8087	0,5279	4,10
AZZA3	103	0,0108	0,5111	4,76	0,5523	0,3693	67,91
B3SA3	103	-0,8724	0,0696	83,50	0,7856	0,5492	0,00
BALM4	103	-0,0355	0,6602	2,50	0,5686	0,5704	99,25
BAZA3	103	0,2445	0,1268	68,93	0,4126	0,1267	65,67
BBAS3	103	-0,3611	0,2143	30,10	0,9925	0,8381	0,00
BBDC3	103	0,0644	0,2466	28,16	0,3189	-0,1838	0,00
BBDC4	103	-0,0038	0,0574	70,87	-0,0192	0,0653	0,00
BBSE3	103	-0,4008	0,1053	77,67	0,7797	0,9042	0,00
BDLL4	103	-0,0535	0,7652	1,47	-0,4013	-0,2979	99,25
BEEF3	103	-0,3153	0,1319	40,78	0,2654	0,2463	0,00

Ticker B3	**N. Prev.**	$\hat{\gamma}$ **Médio**	**p-valor Méd.**	**% Signif.**	$\hat{\alpha}+\hat{\beta}$ **M1**	$\hat{\alpha}+\hat{\beta}$ **M2**	**SVI Zeros**
BEES3	103	0,5025	0,2289	17,48	0,8744	0,7279	9,70
BEES4	103	0,2317	0,6245	5,83	1,0838	1,0107	70,90
BGIP3	103	0,0649	0,6543	1,94	-0,2935	-0,3561	99,25
BGIP4	103	0,2215	0,0868	52,43	1,0162	0,6266	95,90
BHIA3	103	0,5732	0,0198	87,38	0,5121	-0,0080	50,37
BIOM3	103	0,0161	0,1963	49,51	0,2162	0,1398	92,54
BMEB3	103	0,0915	0,1690	51,14	0,8244	0,6754	96,27
BMEB4	103	0,0241	0,0773	68,93	1,0936	1,0860	69,78
BMGB4	103	0,1531	0,1081	57,28	1,0149	0,9665	0,00
BMIN4	103	0,0111	0,1831	58,25	0,9850	0,9168	97,01
BMKS3	103	0,0799	0,2502	13,59	0,7783	0,7475	95,15
BNBR3	103	-0,4712	0,2514	67,35	0,2545	-0,7981	97,76
BOBR4	103	-0,1971	0,1892	47,57	0,6484	0,2372	80,60
BPAC11	103	-0,3323	0,0420	85,44	0,5705	0,5930	0,00
BPAC3	103	-0,1149	0,2478	51,46	0,6196	0,5682	76,12
BPAC5	103	0,0991	0,3460	5,26	0,7966	0,7161	98,13
BRAP3	103	0,1433	0,1709	36,89	0,4618	-0,3475	13,06
BRAP4	103	-0,2378	0,3592	20,39	0,5728	0,1899	0,00
BRAV3	103	-0,1810	0,4610	6,80	0,7240	0,7146	69,03
BRKM3	103	0,0195	0,3688	22,33	0,4270	0,2912	62,69
BRKM5	103	-0,1657	0,4453	13,59	0,2593	0,3169	0,00
BRKM6	103	-0,8867	0,0426	67,96	0,6177	0,4805	99,25
BRSR3	103	-0,1535	0,0411	89,00	0,5116	-0,4856	80,22
BRSR5	103	0,0914	0,6266	0,00	0,5100	0,5027	99,25
BRSR6	103	0,2082	0,2387	27,18	-0,0002	-0,0329	0,00
CAMB3	103	0,0287	0,2003	54,37	0,8060	0,6136	72,76
CAML3	103	0,4744	0,0787	66,02	0,2788	0,0315	0,37
CASH3	103	0,5568	0,0056	98,06	0,7139	0,7167	0,00
CBEE3	103	-0,0540	0,1540	70,69	0,9203	0,9731	98,51
CEAB3	103	-0,1495	0,3677	9,71	0,3117	-0,0633	0,00
CEBR3	103	0,0925	0,1337	45,63	-0,1435	-0,0924	70,52
CEBR5	103	-0,0150	0,7088	5,83	0,2266	0,0142	96,64
CEBR6	103	0,0618	0,3100	58,25	0,1670	-0,1548	64,55
CEDO3	103	-0,0087	0,1947	80,00	0,6070	0,6912	99,63
CEDO4	103	-0,0351	0,4253	7,77	0,8330	0,9400	97,39
CEEB3	103	-13,6478	0,4299	16,18	0,6358	0,3612	96,64
CGAS3	103	0,0954	0,2312	30,10	0,7282	0,3936	94,03
CGAS5	103	-3,1929	0,3123	45,63	0,8418	0,8496	76,87
CGRA3	103	-0,6273	0,0105	96,08	0,8878	0,8263	95,15
CGRA4	103	-0,3647	0,3328	33,01	0,5956	-0,2337	75,37
CLSC4	103	-6,9460	0,5297	16,13	0,3928	0,1115	88,43
CMIG3	103	-0,1862	0,3641	30,10	0,2225	0,5450	2,24
CMIG4	103	-0,5592	0,2735	48,54	0,4945	0,1575	0,00
COCE5	103	0,0819	0,2747	36,89	0,3461	-0,1113	90,67
COGN3	103	0,3810	0,2033	8,74	0,4214	-0,0932	0,00

Ticker B3	**N. Prev.**	$\hat{\gamma}$ **Médio**	**p-valor Méd.**	**% Signif.**	$\hat{\alpha}+\hat{\beta}$ **M1**	$\hat{\alpha}+\hat{\beta}$ **M2**	**SVI Zeros**
CPFE3	103	0,1756	0,3209	3,88	0,6100	0,3050	0,37
CPLE3	103	0,5785	0,1233	66,99	0,6781	0,6763	1,12
CRPG5	103	-0,0311	0,3373	22,33	0,7082	0,5709	75,75
CRPG6	103	-0,3503	0,0797	50,49	0,5003	0,3324	98,88
CSAN3	103	0,0076	0,2714	12,62	0,2813	0,3363	0,00
CSMG3	103	-0,6189	0,0451	80,58	0,3695	0,5485	0,00
CSNA3	103	-0,7143	0,1035	4,85	0,5168	0,6901	0,00
CSUD3	103	0,0854	0,3620	5,83	0,2322	0,1437	50,75
CTKA4	103	0,1726	0,3073	0,00	0,5095	0,4386	98,88
CTSA3	103	-0,1165	0,3205	56,31	-0,4968	-0,2954	91,04
CTSA4	103	-0,1404	0,1197	70,87	0,9056	0,9455	92,16
CURY3	103	-0,0149	0,4125	24,27	0,3381	0,0222	12,31
CVCB3	103	1,1951	0,0598	84,47	0,2200	0,0464	0,00
CYRE3	103	-0,4169	0,1846	42,72	0,1578	0,1284	0,00
DASA3	103	0,0923	0,4072	12,62	0,8432	0,8640	13,43
DEXP3	103	-0,0116	0,2913	24,27	0,3507	0,3180	19,78
DEXP4	103	-0,0642	0,0792	73,33	0,3543	0,3699	99,25
DIRR3	103	-0,0046	0,4430	12,62	0,4112	0,2792	0,75
DMVF3	103	0,0141	0,7174	0,97	0,0179	0,1151	50,00
DOHL4	103	-1,9153	0,6417	25,00	0,5699	0,5140	98,13
EALT3	103	0,2807	0,2530	10,00	0,5877	0,0820	98,51
EALT4	103	0,0173	0,3783	35,92	0,7756	0,7803	73,88
ECOR3	103	-0,1301	0,1968	73,79	-0,3572	0,1751	0,37
EGIE3	103	0,0768	0,4408	33,01	0,7969	0,6304	0,00
EKTR4	103	0,1455	0,2033	16,50	0,2241	-0,0430	98,13
EMAE4	103	0,0500	0,4261	8,74	1,0155	0,9653	81,72
ENEV3	103	-0,2711	0,0941	57,28	0,6747	0,1942	4,48
ENGI11	103	-0,0382	0,4136	17,48	0,5748	0,2366	35,45
ENGI3	103	-0,0346	0,4672	9,71	0,2612	0,0567	61,94
ENGI4	103	0,0084	0,7523	0,00	0,7923	0,6570	76,49
ENJU3	103	-0,2646	0,2975	35,92	0,3699	0,1612	7,46
ENMT3	103	-0,1459	0,1144	81,55	0,4467	0,5505	98,51
EPAR3	103	0,0452	0,2241	47,57	0,9374	0,7174	90,67
EQMA3B	103	0,2956	0,0480	73,74	0,7285	0,5954	99,63
EQPA3	103	-0,2722	0,5066	5,05	0,4151	-0,3222	90,67
EQTL3	103	-0,2870	0,2811	16,50	0,6087	-0,3660	0,00
ESPA3	103	0,0105	0,4796	13,59	0,9911	0,9371	8,96
ESTR4	103	-0,0027	0,1198	73,79	0,7595	0,5451	98,13
ETER3	103	-0,0510	0,1228	75,73	1,0350	0,9694	1,87
EUCA3	103	-0,0139	0,5325	13,59	1,1406	1,1275	95,15
EUCA4	103	-0,2900	0,0421	81,55	0,6763	0,9537	74,63
EVEN3	103	0,0265	0,1865	36,89	0,1460	0,1670	0,00
EZTC3	103	-0,4820	0,1599	59,22	0,1761	0,3694	0,00
FESA3	103	-0,1163	0,2369	45,63	0,7015	0,2843	97,01
FESA4	103	-0,2097	0,2053	55,34	1,0167	0,9409	0,00

Ticker B3	**N. Prev.**	$\hat{\gamma}$ **Médio**	**p-valor Méd.**	**% Signif.**	$\hat{\alpha}+\hat{\beta}$ **M1**	$\hat{\alpha}+\hat{\beta}$ **M2**	**SVI Zeros**
FHER3	103	-1,0198	0,0405	85,44	0,0808	0,3121	41,79
FLRY3	103	-0,4605	0,1406	59,22	0,6806	0,3940	0,00
FRAS3	103	0,1013	0,3852	19,42	0,5871	0,4784	21,64
GEPA3	103	-2,3880	0,3895	42,72	0,7949	0,5614	99,25
GEPA4	103	-2,4223	0,1683	42,86	0,3209	0,1877	97,39
GFSA3	103	0,4855	0,0327	84,47	1,1370	0,9549	0,00
GGBR3	103	-0,0482	0,2857	24,27	0,6160	-0,0488	8,21
GGBR4	103	-0,2431	0,2950	38,84	0,1132	0,1886	0,00
GMAT3	103	-0,4919	0,1715	53,40	0,4836	0,4469	0,37
GOAU3	103	-0,1125	0,2759	21,36	0,4811	0,1258	17,54
GOAU4	103	-0,2311	0,3765	23,30	0,2382	0,1737	0,00
GRND3	103	0,1503	0,3041	28,16	-0,0173	-0,1892	0,00
GSHP3	103	-0,0608	0,0000	100,00	1,0719	1,0873	99,25
HAGA3	103	-0,0708	0,2966	33,01	0,6847	0,7317	96,27
HAGA4	103	0,1122	0,3671	17,48	-0,1775	-0,3184	87,69
HAPV3	103	-0,5079	0,0982	88,35	0,4463	0,1958	0,00
HBOR3	103	-0,3515	0,0695	75,73	0,8718	0,9095	1,49
HBRE3	103	0,1416	0,0568	54,37	0,6624	0,4525	76,87
HBSA3	103	0,1140	0,2173	66,02	0,2231	-0,0042	22,76
HBTS5	103	-0,1365	0,0099	92,06	0,3103	0,0921	98,88
HETA4	103	-0,3982	0,0546	76,70	1,0496	1,0461	98,88
HOOT4	103	-0,0236	0,1828	52,43	0,7465	0,5973	93,66
HYPE3	103	0,3267	0,1954	67,96	0,8748	0,5648	0,00
INEP3	103	-0,0840	0,0668	71,84	0,8261	0,7812	35,07
INEP4	103	0,0158	0,3852	29,13	0,4785	0,2901	47,39
IRBR3	103	0,1132	0,4232	30,10	0,5775	0,5483	0,00
ISAE3	103	-1,2637	0,2945	31,11	0,8688	0,4620	83,58
ISAE4	103	-0,0359	0,7762	0,00	0,6856	0,5823	73,88
ITSA3	103	-0,7199	0,1239	53,40	0,7525	0,0991	0,00
ITSA4	103	-1,6171	0,0757	74,76	0,8571	-0,4596	0,00
ITUB3	103	-0,4963	0,3076	17,48	0,8234	0,2447	0,00
ITUB4	103	-1,3674	0,0404	72,82	0,8106	0,0238	0,00
JFEN3	103	0,0820	0,2870	24,27	0,4131	0,3116	64,93
JHSF3	103	-0,2170	0,1937	51,46	0,4359	0,4871	0,00
JSLG3	103	0,0070	0,5454	17,48	0,7348	0,7657	24,63
KEPL3	103	-0,2554	0,2060	36,89	0,2444	0,1857	2,24
KLBN11	103	-0,6829	0,1507	20,39	0,8788	0,5087	0,00
KLBN3	103	-1,4054	0,0034	100,00	0,3689	0,0055	0,75
KLBN4	103	-0,6806	0,1232	12,62	0,7940	0,2751	0,00
LAVV3	103	0,0544	0,5610	9,71	0,5792	0,4673	11,94
LEVE3	103	-0,0249	0,3631	45,63	0,8024	0,8516	0,00
LIGT3	103	0,3234	0,1828	51,46	0,8801	0,4437	3,73
LJQQ3	103	-0,0986	0,5333	11,65	0,6733	0,6395	0,00
LOGG3	103	-0,0310	0,4226	4,85	0,3160	0,3872	22,76
LOGN3	103	-0,1697	0,0347	85,44	0,6791	0,3454	57,84

Ticker B3	**N. Prev.**	**$\hat{\gamma}$ Médio**	**p-valor Méd.**	**% Signif.**	**$\hat{\alpha}+\hat{\beta}$ M1**	**$\hat{\alpha}+\hat{\beta}$ M2**	**SVI Zeros**
LPSB3	103	-0,0034	0,6091	7,77	-0,0279	-0,0581	52,99
LREN3	103	0,0063	0,7367	1,94	-0,5433	-0,4538	0,00
LUPA3	103	-0,0223	0,1466	27,18	0,9443	0,9487	9,33
LUXM4	103	-0,6338	0,0	0,0	1,1069	0,5353	99,63
LWSA3	103	-0,7086	0,0415	87,38	0,4996	0,5333	0,00
MBRF3	103	0,0190	0,2829	4,17	0,4714	0,4265	90,30
MDIA3	103	-0,0873	0,2107	30,10	0,5845	0,6973	0,75
MDNE3	103	0,0072	0,1203	73,79	0,2461	0,2703	14,93
MEAL3	103	-0,0507	0,3677	48,54	0,8071	0,7737	4,48
MELK3	103	-0,0320	0,3317	44,66	-0,0436	-0,1710	17,91
MGEL4	103	-0,7472	0,0480	73,79	0,8087	0,8111	99,63
MGLU3	103	-0,3248	0,1668	36,89	0,5837	0,5851	0,00
MILS3	103	-0,0049	0,7061	0,97	0,9625	0,9139	10,07
MNDL3	103	0,0215	0,5110	14,00	0,2500	0,1860	98,13
MNPR3	103	0,2274	0,0201	87,38	0,2826	0,2740	95,15
MOVI3	103	-0,7724	0,1267	49,51	0,4141	0,6330	0,00
MRSA3B	103	-7,2608	0,0199	93,14	0,0598	-0,0307	99,63
MRVE3	103	-0,3096	0,2509	49,51	0,7120	0,3800	0,00
MTRE3	103	-0,1834	0,1365	51,46	0,9089	0,8411	5,22
MTSA4	103	-0,1707	0,1206	58,25	0,3422	0,3872	97,01
MULT3	103	0,0869	0,3597	18,45	0,5019	0,5072	2,99
MWET4	103	-0,0931	0,2508	51,46	0,7818	0,7115	94,78
MYPK3	103	0,2091	0,2048	20,39	0,4306	0,1652	0,37
NEOE3	103	-0,2131	0,6297	13,59	0,1402	0,1185	0,37
NEXP3	103	0,0500	0,0889	57,28	1,0476	1,0373	88,06
NGRD3	103	-0,0417	0,3183	17,48	0,9883	0,9187	7,46
NUTR3	103	-0,2684	0,1512	59,22	0,6427	0,7162	98,51
ODPV3	103	0,0061	0,6830	3,88	0,8550	0,6750	6,34
OFSA3	103	-0,0836	0,3523	27,18	-0,2096	-0,0761	92,16
OIBR3	103	1,0820	0,2896	56,31	0,3508	-0,0955	0,00
OIBR4	103	0,0527	0,4466	16,50	0,5998	0,6801	5,22
OSXB3	103	-0,0405	0,4098	25,24	1,1221	0,9625	77,99
PATI3	103	0,2303	0,1180	75,73	0,7641	0,4802	97,39
PATI4	103	0,0334	0,0002	100,00	0,5593	0,5295	99,63
PCAR3	103	-0,7271	0,0021	99,03	-0,0197	-0,2078	0,00
PDGR3	103	0,6180	0,0416	74,76	-0,1998	-0,3532	1,12
PDTC3	103	0,0199	0,6133	15,53	0,4896	0,3781	54,85
PEAB3	103	-0,2682	0,1201	11,11	0,4592	0,5116	99,63
PEAB4	103	-1,5014	0,0430	92,23	0,3842	0,1016	97,76
PETR3	103	-0,8743	0,0635	66,02	0,7201	0,6783	0,00
PETR4	103	-0,4779	0,1877	11,65	0,7582	0,6432	0,00
PFRM3	103	-0,0399	0,4821	0,00	0,6616	0,7051	58,21
PGMN3	103	-0,2486	0,0448	67,96	0,3566	0,2339	1,12
PINE3	103	0,8732	0,0721	82,52	0,8421	0,9822	96,64
PINE4	103	-0,0633	0,5269	23,30	0,8023	0,8636	5,97

Ticker B3	**N. Prev.**	**$\hat{\gamma}$ Médio**	**p-valor Méd.**	**% Signif.**	**$\hat{\alpha}+\hat{\beta}$ M1**	**$\hat{\alpha}+\hat{\beta}$ M2**	**SVI Zeros**
PLAS3	103	-0,1536	0,3257	18,75	0,8189	0,7399	98,51
PLPL3	103	-0,0038	0,3587	21,36	0,3984	0,3925	4,10
PMAM3	103	0,2941	0,0232	95,15	0,6764	0,5267	3,36
PNVL3	103	-0,0275	0,2843	24,27	0,7691	0,5474	11,57
POMO3	103	-0,4097	0,0409	89,32	0,2565	0,4442	1,49
POMO4	103	0,5569	0,1021	79,61	0,4341	0,0781	0,00
POSI3	103	-0,1550	0,4002	33,98	0,9626	0,7341	0,00
PPLA11	103	-0,0287	0,4499	0,00	0,2016	0,2170	99,25
PRIO3	103	-0,3885	0,2216	29,13	0,6844	0,3455	0,00
PRNR3	103	-0,0410	0,1143	48,54	0,2484	-0,1236	37,31
PSSA3	103	-0,2464	0,4464	3,88	0,0605	0,0610	0,00
PTBL3	103	-0,0471	0,3638	21,36	-0,0462	-0,2662	3,73
PTNT3	103	0,0520	0,6297	6,80	0,7282	0,7124	97,01
PTNT4	103	0,2167	0,3020	6,10	0,4931	0,3891	97,39
QUAL3	103	-0,6550	0,0907	66,02	0,1707	0,1267	0,00
RADL3	103	0,2309	0,2342	22,33	-0,1612	-0,3442	0,00
RAIL3	103	-0,1861	0,2783	38,84	0,4098	0,1215	0,00
RANI3	103	-0,0061	0,2147	53,40	0,8010	0,3106	0,00
RAPT3	103	0,0133	0,2118	57,28	0,4320	0,3553	86,57
RAPT4	103	0,1192	0,4216	8,74	0,3775	0,3329	0,75
RCSL3	103	-0,0265	0,4156	38,84	1,0683	1,1083	31,72
RCSL4	103	-0,2828	0,0033	99,03	0,8270	1,0925	19,78
RDOR3	103	-0,3174	0,1359	36,89	0,7464	0,7879	0,00
REDE3	103	-0,5740	0,0245	92,08	0,8541	0,5641	92,91
RENT3	103	-0,4192	0,1686	69,90	0,7226	0,7603	0,00
RNEW3	103	0,0209	0,4537	30,10	0,4118	-0,0056	66,42
RNEW4	103	0,0073	0,3327	31,07	0,5616	0,4216	44,78
ROMI3	103	-0,3569	0,1058	52,43	0,5472	0,0868	0,00
RPAD3	103	0,1958	0,2744	67,07	0,0373	-0,4541	99,25
RPAD5	103	-3,3268	0,0431	63,24	0,8767	0,7033	99,63
RPAD6	103	-5,3567	0,0356	87,38	-0,2882	0,0624	99,63
RPMG3	103	-0,1361	0,3121	40,78	0,9637	1,0440	80,97
RSID3	103	0,0195	0,3703	6,80	0,5741	0,5453	86,57
RSUL4	103	-0,1290	0,2777	32,04	0,6726	0,4522	96,27
SANB11	103	-0,0890	0,6913	1,94	0,3300	0,2149	0,00
SANB3	103	-0,2107	0,4994	4,85	0,2932	-0,2393	0,75
SANB4	103	-0,1503	0,5613	18,45	0,3296	0,0629	1,12
SAPR11	103	-0,5617	0,1198	44,66	0,7342	0,5796	0,00
SAPR3	103	0,0384	0,1264	38,84	0,5397	0,3223	0,00
SAPR4	103	-0,2102	0,6344	5,83	0,7057	-0,1029	0,00
SBFG3	103	0,1393	0,0902	71,84	-0,1158	-0,0459	5,60
SBSP3	103	0,1520	0,3041	9,71	0,7871	0,5794	0,37
SCAR3	103	0,2057	0,2654	23,30	0,7028	0,7071	88,43
SEER3	103	0,0910	0,1740	9,71	0,7762	0,8051	13,43
SEQL3	103	-0,1906	0,3791	14,56	0,1544	0,1802	2,24

Ticker B3	**N. Prev.**	$\hat{\gamma}$ **Médio**	**p-valor Méd.**	**% Signif.**	$\hat{\alpha}+\hat{\beta}$ **M1**	$\hat{\alpha}+\hat{\beta}$ **M2**	**SVI Zeros**
SHOW3	103	-0,0060	0,4628	14,56	0,8501	0,7990	30,97
SHUL4	103	-0,1196	0,3016	13,59	0,6714	0,7590	2,24
SIMH3	103	-0,3502	0,0667	86,41	0,3581	0,5035	0,00
SLCE3	103	-0,1164	0,2796	38,84	0,7429	0,3998	0,00
SMTO3	103	-0,0742	0,3030	23,30	0,7683	0,8004	1,87
SUZB3	103	-0,7035	0,0724	68,93	0,5490	0,3524	0,00
SYNE3	103	1,0901	0,0120	88,35	-0,1370	-0,2210	15,67
TAEE11	103	0,7531	0,0793	59,22	0,8740	0,6727	0,00
TAEE3	103	0,7272	0,0390	92,23	0,7629	0,6918	0,00
TAEE4	103	0,6202	0,0873	60,19	0,9568	0,6223	0,00
TASA3	103	0,0381	0,1499	53,40	0,7368	0,8186	44,03
TASA4	103	0,4205	0,1364	56,31	0,4431	0,4713	0,00
TCSA3	103	-0,1471	0,1908	37,86	0,3508	0,2868	32,09
TECN3	103	0,0431	0,4419	11,65	0,9174	0,8969	23,88
TELB3	103	-0,0771	0,0375	83,50	0,5075	0,6436	91,42
TELB4	103	-0,0651	0,1068	68,93	0,3995	0,1709	77,61
TEND3	103	-0,1332	0,2693	33,98	0,5804	0,5124	3,36
TFCO4	103	0,0572	0,2441	42,72	0,2620	0,2189	90,67
TGMA3	103	0,0166	0,1130	60,19	0,6793	0,4155	54,85
TIMS3	103	0,3240	0,2979	10,68	0,3392	0,2984	0,00
TKNO4	103	0,3270	0,0457	64,71	0,1320	-0,0595	99,63
TOTS3	103	0,0525	0,3476	28,16	0,0711	-0,0010	0,00
TPIS3	103	0,1656	0,1077	66,02	0,7781	0,0900	36,57
TRIS3	103	-0,0564	0,4951	2,91	0,9802	0,9838	2,24
TUPY3	103	-0,2259	0,2845	21,36	0,2236	0,4638	0,00
UCAS3	103	0,0068	0,4147	20,39	0,9575	0,9037	85,45
UGPA3	103	-0,2413	0,0614	69,90	0,6759	0,4704	0,37
UNIP3	103	0,0818	0,1799	64,08	0,6045	0,5888	28,73
UNIP5	103	-0,0013	0,4474	15,53	0,7601	0,7987	92,16
UNIP6	103	0,2373	0,3310	32,04	0,4656	0,3453	0,00
USIM5	103	-0,1778	0,2433	43,69	0,8521	0,6815	0,00
VALE3	103	-0,6753	0,1838	38,84	0,6119	0,3976	0,00
VAMO3	103	-0,2289	0,1677	45,63	0,5210	0,3429	0,37
VBBR3	103	0,1748	0,3998	8,74	0,5093	0,4934	13,81
VIVA3	103	-0,4205	0,1588	55,34	0,2689	-0,1588	0,75
VIVR3	103	1,5965	0,0127	94,17	0,3152	-0,2515	6,72
VIVT3	103	-0,0610	0,6283	1,94	0,0997	-0,0284	0,00
VLID3	103	0,1827	0,0350	89,32	0,0555	-0,2789	21,27
VULC3	103	-0,1716	0,1420	68,93	0,6573	-0,0851	3,36
WEGE3	103	0,0354	0,4203	29,13	0,3541	0,4036	0,00
WHRL3	103	-3,4724	0,4114	8,47	0,3278	0,1712	98,51
WHRL4	103	-2,3303	0,4048	32,95	0,8933	0,4760	88,06
WIZC3	103	0,0765	0,3015	13,59	0,7071	0,5222	39,18
WLMM4	103	-0,0723	0,6118	17,48	0,3145	0,3494	98,13
YDUQ3	103	-0,0302	0,5153	6,80	0,8617	0,8489	0,00

Fonte: elaboração própria (2026).

REFERÊNCIAS

ANDERSEN, Torben G.; BOLLERSLEV, Tim. Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. **International Economic Review**, [Economics Department of the University of Pennsylvania, Wiley, Institute of Social e Economic Research, Osaka University], v. 39, n. 4, p. 885–905, 1998. ISSN 00206598, 14682354.

ANDERSEN, Torben G.; BOLLERSLEV, Tim et al. Modeling and Forecasting Realized Volatility. **Econometrica**, v. 71, n. 2, p. 579–625, 2003. DOI: 10.1111/1468-0262.00418.

ANDREI, Daniel; HASLER, Michael. Investor Attention and Stock Market Volatility. **The Review of Financial Studies**, v. 28, n. 1, p. 33–72, 2015. DOI: 10.1093/rfs/hhu059.

B3 S.A. – BRASIL, BOLSA, BALCÃO. **Book Pessoa Física – 1º Trimestre de 2025**. Versão portuguesa. São Paulo, 2025. Disponível em: https://www.b3.com.br/data/files/9E/70/C2/F1/2C1179106B8BCB69AC094EA8/Book%20PF%201Tri25.pdf. Acesso em: 1 jun. 2026.

BARBER, Brad M.; ODEAN, Terrance. All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. **The Review of Financial Studies**, v. 21, n. 2, p. 785–818, 2008. DOI: 10.1093/rfs/hhm079.

BOLLEN, Johan; MAO, Huina; ZENG, Xiaojun. Twitter mood predicts the stock market. **Journal of Computational Science**, v. 2, n. 1, p. 1–8, 2011. DOI: 10.1016/j.jocs.2010.12.007.

BOLLERSLEV, Tim. Generalized autoregressive conditional heteroskedasticity. **Journal of Econometrics**, v. 31, n. 3, p. 307–327, 1986. ISSN 0304-4076. DOI: 10.1016/0304-4076(86)90063-1.

BORDINO, Ilaria et al. Web Search Queries Can Predict Stock Market Volumes. **PLoS ONE**, v. 7, n. 7, e40014, 2012. DOI: 10.1371/journal.pone.0040014.

DA, ZHI; ENGELBERG, JOSEPH; GAO, PENGJIE. In Search of Attention. **The Journal of Finance**, v. 66, n. 5, p. 1461–1499, 2011. DOI: 10.1111/j.1540-6261.2011.01679.x.

DICKEY, David A.; FULLER, Wayne A. Distribution of the Estimators for Autoregressive Time Series With a Unit Root. **Journal of the American Statistical Association**, v. 74, p. 427–431, 1979. DOI: 10.2307/2286348.

FAMA, Eugene F. Efficient Capital Markets: A Review of Theory and Empirical Work. **The Journal of Finance**, v. 25, n. 2, p. 383–417, 1970. DOI: 10.2307/2325486.

FAMA, Eugene F. The Behavior of Stock-Market Prices. **The Journal of Business**, v. 38, n. 1, p. 34–105, 1965.

GERVAIS, Simon; KANIEL, Ron; MINGELGRIN, Dan H. The High-Volume Return Premium. **The Journal of Finance**, v. 56, n. 3, p. 877–919, 2001. DOI: 10.1111/0022-1082.00349.

GHALANOS, Alexios. **rugarch: Univariate GARCH Models**. [S.l.], 2024. R package. Disponível em: https://cran.r-project.org/package=rugarch.

GOOGLE. **FAQ about Google Trends data**. Disponível em: https://support.google.com/trends/answer/4365533?hl=en-GB&ref_topic=6248052. Acesso em: 4 maio 2026.

GUJARATI, Damodar N; PORTER, Dawn C. **Econometria Básica**. 5. ed. Porto Alegre: AMGH Editora, 2011. Tradução de Basic Econometrics, 5th ed. ISBN 9788580550511.

HANSEN, Peter R.; LUNDE, Asger; NASON, James M. The Model Confidence Set. **Econometrica**, v. 79, n. 2, p. 453–497, 2011. DOI: 10.3982/ECTA5771.

HIRSHLEIFER, David; TEOH, Siew Hong. Limited attention, information disclosure, and financial reporting. **Journal of Accounting and Economics**, v. 36, n. 1-3, p. 337–386, 2003. DOI: 10.1016/j.jacceco.2003.10.002.

JANG, Jaehee. **YouTube, View count, Stock market: Investor attention and stock returns**. Working paper (dissertação de mestrado), Hanyang University, Seul. 2023. Disponível em: https://ssrn.com/abstract=4669452. Acesso em: 28 maio 2026.

KAHNEMAN, Daniel. **Attention and Effort**. Englewood Cliffs, NJ: Prentice-Hall, 1973.

MAVRAGANI, Amaryllis; OCHOA, Gabriela. Google Trends in Infodemiology and Infoveillance: Methodology Framework. **JMIR Public Health Surveill**, v. 5, n. 2, e13439, maio 2019. ISSN 2369-2960. DOI: 10.2196/13439. Disponível em: http://publichealth.jmir.org/2019/2/e13439/.

MERTON, Robert C. A Simple Model of Capital Market Equilibrium with Incomplete Information. **The Journal of Finance**, v. 42, n. 3, p. 483–510, 1987. DOI: 10.1111/j.1540-6261.1987.tb04565.x.

NELSON, Daniel B. Conditional Heteroskedasticity in Asset Returns: A New Approach. **Econometrica**, v. 59, n. 2, p. 347–370, 1991. DOI: 10.2307/2938260.

PATTON, Andrew J. Volatility forecast comparison using imperfect volatility proxies. **Journal of Econometrics**, v. 160, n. 1, p. 246–256, jan. 2011.

PENG, Lin; XIONG, Wei. Investor attention, overconfidence and category learning. **Journal of Financial Economics**, v. 80, n. 3, p. 563–602, 2006. DOI: 10.1016/j.jfineco.2005.05.003.

PERLIN, Marcelo 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, 2017. DOI: 10.1002/for.2446.

PREIS, Tobias; MOAT, Helen Susannah; STANLEY, H. Eugene. Quantifying Trading Behavior in Financial Markets Using Google Trends. **Scientific Reports**, v. 3, p. 1684, 2013. DOI: 10.1038/srep01684.

SIMS, Christopher A. Implications of rational inattention. **Journal of Monetary Economics**, v. 50, n. 3, p. 665–690, 2003. DOI: 10.1016/S0304-3932(03)00029-1.

STATCOUNTER GLOBAL STATS. **Search Engine Market Share Brazil**. [S.l.*:* s.n.], 2025. https://gs.statcounter.com/search-engine-market-share/all/brazil. Acesso em: 1 jun. 2026.

STOCK, James H.; WATSON, Mark W. **Econometria**. São Paulo: Pearson, 2004.

VOZLYUBLENNAIA, Nadia. Investor attention, index performance, and return predictability. **Journal of Banking & Finance**, v. 41, p. 17–35, 2014. DOI: 10.1016/j.jbankfin.2013.12.010.