ISSN 1808-057X
DOI: 10.1590/1808-057x20231635.en

Original Article

# The effect of investor attention on the efficiency of the Brazilian stock market

**Marcelo Guzella**[1]
https://orcid.org/0000-0001-5243-8897
Email: marcelo.guzella@professores.ibmec.edu.br

**F. Henrique Castro**[2]
https://orcid.org/0000-0001-7456-2354
Email: henrique.castro@fgv.br

**Verônica de Fátima Santana**[3]
https://orcid.org/0000-0002-9105-7488
Email: veronica.santana@fecap.br

[1] Ibmec BH, Departamento de Economia, Belo Horizonte, MG, Brazil
[2] Fundação Getulio Vargas, Escola de Economia de São Paulo, São Paulo, SP, Brazil
[3] Fundação Escola de Comércio Álvares Penteado, São Paulo, SP, Brazil

Received on 07.21.2022 – Desk acceptance on 08.01.2022 – 2nd version approved on 05.29.2023
Editor-in-Chief: approved by Fábio Frezatti, published by Andson Braga de Aguiar
Associate Editor: Fabio Gallo Garcia

## ABSTRACT

The objective of this article was to evaluate the relationship between investor attention, both professional and non-professional, and the efficiency of the Brazilian stock market, as measured by the predictability of daily returns. The role of investor attention in the capital market is a controversial issue, since some studies show that it is capable of inducing efficiency, while others point out that it essentially contributes to greater volatility due to behavioral biases. This study contributes to the behavioral finance literature by pioneering the analysis of the effect of investor attention on Brazilian stock prices. Moreover, we found no previous studies that compare the effect of investor attention on market efficiency across investor classes, as measured by access to different financial information providers. The analysis of investor attention, as measured by the volume of searches on the internet, is relevant given the growing abundance of data and digital inclusion. The evidence of an effect of attention on market efficiency contributes to the critique of this classical hypothesis. The methodology consisted of estimating autoregressive models using daily data from 2018 to 2021 on search volume, returns, and transaction volume for Brazilian stocks, including the interaction between returns and search volume, to assess the influence of attention on price dynamics. In this work, we find that in the Brazilian stock market, investor attention contributes to greater market efficiency, as measured by lower predictability of returns, but the most pronounced effect comes from the attention of non-professional investors. Understanding how attention affects the incorporation of information into prices contributes to a critique of the market efficiency hypothesis, as well as allowing for profits to be made by exploiting opportunities based on the level of attention of a particular class of investors.

**Keywords:** behavioral finance, investor attention, volume of internet searches, predictability of returns.

**Correspondence address**

**Marcelo Guzella**
Ibmec BH, Departamento de Economia
Rua Rio Grande do Norte, 300 – CEP: 30130-130
Funcionários – Belo Horizonte – MG – Brazil

This is a bilingual text. This article was originally written in Portuguese and published under the DOI https://doi.org/10.1590/1808-057x20231635.pt

This article stems from a doctoral dissertation defended by the author, Marcelo Guzella, in 2020.

Paper presented at the XIX Brazilian Meeting on Finance, July 2019, Rio de Janeiro, RJ, Brazil, and at the 6th Brazilian Meeting on Economics and Behavioral Finance, August 2019, São Paulo, SP, Brazil.

## 1. INTRODUCTION

The hypothesis of immediate incorporation of relevant information into asset prices is one of the best known in finance. The theoretical justification is that incentives, the aggregation of information among investors, and the actions of arbitrageurs ensure this incorporation. However, there are phenomena that challenge these arguments, such as the autocorrelation of returns and the effects of reversals and momentum. Theoretical models have attempted to explain these anomalies, such as the work of Amihud (2002) and Barberis et al. (1998).

It is reasonable to imagine that these anomalies could also result from fluctuations in aggregate levels of attention. Attention is a scarce resource that varies over time. Work in the field of psychology documents that people can only digest a subset of the available information due to time constraints, which can lead to over- or under-reactions. This is the case in Sims' (2005) model, which incorporates the idea that most people can easily find much more information relevant to their decision-making than they actually use to decide. This so-called irrational inattention is based on the premise that agents have to select the relevant information from what is available.

Several studies have addressed this issue in recent years. Falkinger's (2008) model, which states that companies compete first for people's attention and then compete for their propensity to consume, helps explain documented phenomena such as home bias (Mondria & Wu, 2010; Van Nieuwerburgh & Veldkamp, 2007) and the equity premium puzzle (Gabaix & Laibson, 2005).

In this context, we empirically investigate the impact of limited information processing capacity on stock prices. Attention to so much information is crucial for the behavior of markets, which justifies understanding the effect transmitted. On the one hand, excessive attention from non-professional investors can create additional noise and volatility (Barber & Odean, 2008; Da et al., 2011). On the other hand, more attention can mean that the market absorbs more information, thereby increasing its efficiency. This conjecture is called the information discovery hypothesis (Fang & Peress, 2009; Grullon et al., 2004). It states that more information absorbed and reflected in prices essentially makes markets less predictable and therefore more efficient.

Several previous studies have developed analyses to better understand the effect of investor attention on financial markets. These include Andrei and Hasler's (2015) model on the role of attention and uncertainty in volatility and risk premiums, and Zhang et al.'s (2013) evaluation of abnormal returns. These studies are important references for examining the role of attention in the context of the Brazilian capital market.

Our work contributes to the behavioral finance literature by pioneering the analysis of the effect of investor attention on the information transmission process to Brazilian stock prices. In addition, we compare the effect on market efficiency of having access to different financial information providers, one being Google for non-professional investors, and the other Bloomberg for professionals. We found no previous studies comparing this type of segregation, which is why we addressed this research problem. In addition to exploring the potential of measuring attention based on internet search volumes, the results contribute to critiquing the market efficiency hypothesis, identifying investor profiles, and gaining a better understanding of market predictability.

In this study, we sought to answer whether the attention of professional and non-professional investors can affect the predictability of the Brazilian stock market. Our findings can also help companies that want to be more visible to investors. Understanding how attention affects the incorporation of information into prices contributes to better management of press releases and publicity. In addition, other market participants (e.g., arbitrageurs) can benefit from their predictability by evaluating the level of attention of a certain class of investors.

The article is structured as follows: the next section presents the theory underlying the study, as well as previous work on related topics. We then describe the methodology used. Finally, we present the results, robustness tests, and concluding remarks.

## 2. LITERATURE REVIEW

Commercial transactions take place between individuals because of heterogeneities in preferences, wealth, or beliefs. Focusing on differences in beliefs, Grossman and Stiglitz (1980) presented a set of conjectures describing the market as a system in which prices, through buy and sell orders, play the role of transferring information from the more informed to the less informed. This system cannot always be in equilibrium, as this would eliminate any possibility of arbitrageurs making profits from their costly activities. For this reason, they theorize that prices only partially reflect the information of arbitrageurs. In this way, the costs of obtaining information are offset.

The seminal work of Grossman and Stiglitz (1980) provides the basis for a paradigm that challenges the idea that prices immediately reflect all available information. In a process of information incorporation based on fluctuations in the number of informed and their expected utility, the level of attention of agents must play an essential role.

Merton (1987) also criticizes the premises of the classical theory, such as the idea that companies can raise immediate and sufficient capital for investments and the little importance given to financial intermediaries. The author asks whether progress in identifying empirical anomalies will lead to the abandonment of the rational behavior paradigm or whether this can be resolved within the traditional framework. A market equilibrium model with heterogeneity in the absorption of information by investors is proposed. The model assumes that an investor is informed about an asset if they know its parameters (expected return, physical investment rate, and production technology constants). The basic condition for an investor to consider an asset in their portfolio is to know about it. Merton (1987) argues that, in equilibrium, the market value of a given firm will always be lower with incomplete information, with an effect similar to an additional discount rate. The larger the investor base, the larger the difference. The additional expected return is proportional to the hidden cost of this incomplete information. The market portfolio is not mean-variance efficient in this incomplete information model.

Sims (2005) examines how the level of attention contributes to the information absorption process and finds that the inertia in the response of economic agents to the availability of external information is due to limited attentional capacity. This inertia is typically treated in economic models by means of adjustment costs or information or implementation lags. Work addressing these issues typically has limitations, such as the assumption that prices are observed without error. This is equivalent to assuming unlimited information processing capacity.

The question is now whether progress in identifying empirical anomalies will lead to a complete overcoming of the rational behavior paradigm, or whether this issue will be resolved based on the traditional model. Grullon et al. (2004) analyzed the effects of this gradual process of incorporating information into prices, assuming that the impact of a firm's visibility among its investors has an impact on the capital market. Based on this hypothesis, they tested whether the amount spent by the firm on advertising affects the liquidity of its shares and its investor base. Their findings, which are robust to different methodologies, show that more visible companies have more liquid shares and a larger number of individual and professional investors in their base.

Using data from 1993 to 2002 for stocks listed on the New York Stock Exchange (NYSE), Fang and Peress (2009) found a negative relationship between media coverage and returns. Media coverage was determined by counting articles in mass-circulation print media. Dividing stocks into tertiles of coverage, they found a 4.8% annual return differential between groups of low and high coverage stocks (equally weighted portfolios). This premium for lack of coverage remains when controlling for various characteristics, such as size.

One of the first studies to empirically investigate the influence of attention on market efficiency using internet data was that of Vozlyublennaia (2014). In a first approach, the author found that the attention to an index has a significant short-term effect on its return. This is consistent with the hypothesis that undiscovered information can be perceived by investors as indicating a higher (lower) return in the future, which would then lead to an increase (decrease) in asset returns. This evidence differs from Barber and Odean's (2008) notion that an increase in attention would generate buying pressure and, consequently, an increase in prices. On the other hand, the impact of returns on attention was long term.

Vozlyublennaia (2014) assessed the dynamics between market variables (return and volatility) and investor attention for six different asset indices: stocks (Dow Jones Industrial, S&P 500, and Nasdaq), bonds (Chicago Board Options Exchange 10-Year Treasury Note Yield

Index), commodities (West Texas Intermediate crude oil index), and gold (Chicago Board Options Exchange Gold). Attention was measured by the volume of searches for each index on Google. The data were obtained on a weekly basis from January 2004 to December 2012.

The effect of attention on the predictability of returns was assessed by including lagged interaction terms of returns and attention in autoregressive models. The results showed significant effects on these interactions and, more importantly, inverse effects to the autocorrelation effects. That is, attention reduces the autocorrelations of returns. In other words, attention shocks reduce the predictability of returns and thus increase market efficiency. The regressions were controlled for possible changes in investment opportunities using macroeconomic variables.

The results showed that the interaction terms indicated the impact of attention on reducing predictability in the Dow, S&P 500, and oil indices. The commodity and bond indices did not show a significant relationship. For the Nasdaq index, the autocorrelations and interactions had the same sign. Attention shocks also showed a reduction in the predictability of short-term volatility in the Dow, Nasdaq, gold, and oil indices.

Tantaopas et al. (2016) also investigated the relationship between attention and market efficiency by examining the impact of the number of internet searches on market variables (return, volatility, and trading volume). The authors looked at developed and developing markets in the Asia-Pacific region, which are relatively underexplored for this type of study. According to the authors, the information discovery hypothesis follows a cyclical process that begins with an abnormal return due to the market's reaction to some corporate event. If this event is enough to attract the attention of some investors, they will search for information about the company's stock. If they find good (bad) news, they may make a decision to buy (sell) the stock, which will result in an increase (decrease) in its price. According to them, positive price pressure due to excessive attention does not explain price declines and does not apply in a reality where short selling is allowed.

The authors obtained weekly data on Google search volumes from January 2004 to December 2014. They calculated indices that represented a significant share of the total market, one per country. Market efficiency was measured by the predictability of returns, volatility, and trading volume. As in Vozlyublennaia's (2014) study, interactions between the lagged components of these series and those of the attention series were used. Using a vector autoregressive (VAR) specification, they found an increase in market efficiency in six countries: Hong Kong, India, Malaysia, Japan, Korea, and Singapore. In the latter three, the presence of an attention unit contributed to a reduction in the predictability of returns by 34.1%, 35.8%, and 36.8%, respectively.

In terms of volatility, internet searches contributed to efficiency gains in eight of the 10 markets (only New Zealand and Thailand did not follow this pattern). In China, the reduction in predictability was 56.1%. In general, the effect of increased efficiency was stronger in developed markets than in developing ones. In the case of volatility, the opposite effect was observed. The ability of attention to reduce the predictability of trading volume was not found to be as significant. This effect was found only in Singapore. The authors also concluded that most of the causal relationships were one-directional: changes in returns, volatility, and trading volume determined changes in attention. In addition, the sign of the relationship depends on the time since the impact, which is short for both developed and developing countries. These studies showed convergent results in assessing the effect of attention on the efficiency of financial markets.

This study breaks new ground by first examining individual Brazilian stocks on a daily basis. It also looks at two different classes of investors: professionals and individuals. This approach helps to test whether this pattern holds in this important developing market. In line with the findings of previous literature, we expect higher levels of attention to reduce the predictability of returns, i.e.:

> $H_1$: increased investor attention promotes greater market efficiency.

Lux and Marchesi (1999) found a relationship between the attention of certain investors and non-fundamental volatility in the markets. Barber and Odean (2008) highlighted the role of non-professional investors in the occurrence of abnormal returns. As this class of investors is more susceptible to behavioral biases (Cronqvist & Siegel, 2014), their strength may affect the proper transmission of information into prices. These biases include salience, anchoring, sunk costs, status quo, representativeness, endowment effect, loss aversion, conjunction fallacy, and overconfidence (Burton & Shah, 2013).

Thus, assuming that our first hypothesis is true, and considering that the excessive attention of non-professional investors is more recognized for generating additional noise and volatility, we present the second hypothesis:

> $H_2$: the attention of professional investors contributes more to market efficiency.

We evaluated the difference between the effects of the two different classes of investors by comparing, in each case and for the same point in time, the magnitudes of (i) the coefficient indicating the predictability of returns and (ii) the coefficient indicating the relationship between returns and the attention of that class. Since these two coefficients have opposite signs, the closer their magnitudes are, the greater the attenuation effect on predictability by attention.

The following sections describe in more detail how and with what information the analysis of these hypotheses is approached.

## 3. METHODOLOGY

In this section, we describe how the hypotheses of attention-induced reductions in predictability were tested. We present details on the sources and properties of the financial and search volume information, as well as the empirical models and expected relationships.

### 3.1 Sample

We chose to analyze the universe of individual stocks because, although the literature on the influence of internet activity on the behavior of indices is scarce, it is even scarcer for individual stocks. Furthermore, the measure of professional investor attention used in this study is only available for this type of asset. Therefore, we chose to form our sample from the 92 stocks that were part of the Bovespa index (Ibovespa) during the second four months of 2022.

The Ibovespa, the most popular Brazilian stock index, was created in 1968 and is a benchmark for investors around the world. The stocks that make up the index represent approximately 85% of the volume traded on the Brazilian capital market (B3 S.A. – Brasil, Bolsa, Balcão [B3], 2015). The portfolio of stocks to be included in the Ibovespa is redefined every 4 months. The criteria for this process mainly take into account trading volume and participation in trading sessions. The assets that make up the Ibovespa are weighted by their respective market values. The asset selection criteria are described by Castro et al. (2019).

Our analysis includes indicators of attention from both professional and non-professional investors. For the attention of non-professional investors, we obtained series of search volume on Google. Google is the most popular search engine in Brazil and in most countries in the world (Statista, 2014). Because it is not a specific tool for financial analysis and because it is free, it is usually used by less sophisticated investors. Da et al. (2011) showed that Google searches capture the attention of non-professional investors by analyzing their correlation with (i) stock suggestions in programs for the general public and (ii) the volume of buy and sell orders from retail investors.

Thus, we obtained from Google Trends daily index series reflecting the search volume of each of the 92 stocks that made up this index in the second 4 months of 2022. To extract the index series representing the volume of searches on Google for each individual stock, we used the exact ticker of each stock as a keyword. This strategy avoids the influence of searches made for purposes other than stock market trading, such as consumption and job opportunities. In addition, only searches made in Brazil were filtered out to avoid time zone effects. Due to the limitations of Google Trends (daily series can be collected for a maximum of 9 months and weekly series for a maximum of 5 years), we first obtained the daily series for 9-month periods separately and then concatenated them using the method proposed by Johansson (2014), starting from the weekly series for the period from January 2018 to December 2022. The search series were then normalized to have a mean of 0 and a unit standard deviation (SD) to facilitate the interpretation of the results.

For professional attention, we extracted the Bloomberg search volume series for each individual stock that was part of the Ibovespa in the second 4 months of 2022 to carry out the analysis. Bloomberg is a paid provider of financial information that has a specific and more complete interface and is widely used by professional and institutional investors, such as portfolio managers and buy-side and sell-side analysts. Ben-Rephael et al. (2017) noted that approximately 80% of Bloomberg users work in the financial sector, including banks and asset managers. Unlike Google searches, Bloomberg searches have a significantly higher correlation with the volume of institutional investor transactions than with total volume. We used the same variable as Ben-Rephael et al. (2017): News Heat – Daily Max Readership, which combines the number of times each article was read by users with the number of times information about a particular stock was searched. The Bloomberg series were exported to Excel using the appropriate add-in, and were also normalized to have a mean of 0 and unit SD.

For the daily series of the two types of attention measures, we only considered days when there was a trading session. As expected, we found several days with no information available, especially at the beginning of the sample period, since some stocks are new. Therefore, the final sample included only 34 of the 92 stocks due to insufficient data. Stocks that did not have data for full quarters throughout the sample and that, after this exclusion, had less than 50% search volume (SV) data available from either Google (GSV) or Bloomberg (BSV) were excluded. Table 1 shows the companies in the final sample and their respective stocks grouped by segment. We have bolded the stocks with the largest representation in the index, which together account for more than 32%.

After obtaining the attention indicators, we obtained the daily log returns adjusted for dividends and turnover for each stock in the sample. The series were obtained from Economatica®.

**Table 1**
*Sample shares*

| Sector | Company | Ticker | % in the Ibovespa |
|---|---|---|---|
| Water and sanitation | Cia de Saneamento Básico de São Paulo | SBSP3 | 0.75 |
| Processed foods | BRF | BRFS3 | 0.73 |
| | JBS | JBSS3 | 2.43 |
| | Marfrig | MRFG3 | 0.32 |
| | Minerva | BEEF3 | 0.17 |
| Banking | Banco Bradesco | BBDC3 | 1.12 |
| | | BBDC4 | 4.61 |
| | Banco do Brasil | BBAS3 | 2.34 |
| | **Itaú Unibanco** | **ITUB4** | **5.66** |
| Beverages | AmBev | ABEV3 | 3.16 |
| Electricity | Centrais Elétricas Brasileiras | ELET3 | 0.72 |
| | | ELET6 | 0.48 |
| | Cia Energética de Minas Gerais | CMIG4 | 0.78 |
| | CPFL Energia | CPFE3 | 0.31 |
| | Eneva | ENEV3 | 0.86 |
| Engineering and construction | MRV Engenharia | MRVE3 | 0.15 |
| Real estate operations | BR Malls Participações | BRML3 | 0.39 |
| Aircraft manufacturing | Embraer | EMBR3 | 0.52 |
| Pharmaceuticals | RaiaDrogasil | RADL3 | 1.11 |
| Metallurgy and mining | Companhia Siderúrgica Nacional | CSNA3 | 0.65 |
| | Gerdau | GGBR4 | 1.52 |
| | Usiminas | USIM5 | 0.29 |
| | **Vale** | **VALE3** | **15.58** |
| Oil, gas, and biofuels | Cosan | CSAN3 | 1.22 |
| | Petro Rio | PRIO3 | 1.11 |
| | **Petrobras** | **PETR3** | **4.49** |
| | | **PETR4** | **6.86** |
| Chemicals | Braskem | BRKM5 | 0.53 |
| Health care | Qualicorp | QUAL3 | 0.18 |
| Insurance | BB Seguridade | BBSE3 | 0.85 |
| Financial services | Cielo | CIEL3 | 0.19 |
| Telecommunications | Telefônica Brasil | VIVT3 | 1.09 |
| Transportation | CCR | CCRO3 | 0.69 |
| Air transportation | Gol Linhas Aéreas Inteligentes | GOLL4 | 0.13 |

**Note:** *Some companies have more than one class of shares with different rights. Representativeness in the market value of the Bovespa index (Ibovespa) refers to the position on June 23, 2022. We have bolded the stocks with the highest representation in the index.*
**Source:** *Prepared by the authors.*

## 3.2 Model

For this test, we adapted the method used by Tantaopas et al. (2016) and Vozlyublennaia (2014). When the market is informationally efficient, there is no serial dependence in asset returns (Fama, 1965). Returns with serial dependence are at least partially predictable. Therefore, more market efficiency implies less predictability of returns.

Tantaopas et al. (2016) and Vozlyublennaia (2014) estimated a VAR of the return and SV series, including an interaction term between the return and SV lags, to test how levels of attention alter the autocorrelation of returns and thus their degree of predictability. While Tantaopas et al. (2016) and Vozlyublennaia (2014) worked with aggregate market indices, our analysis has a panel format, as we analyzed several stocks (N) over several time periods (T). Usually, panel data estimations depend on asymptotic assumptions that are valid for short panels, i.e., when N > T and N tends to infinity. In this study, however, we have N = 34 and T = 940, forming a long panel in which inferences must be made based on the assumption that T tends to infinity, i.e., time series characteristics must be incorporated into the model. According to Cameron and Trivedi (2005), long panels can be estimated by incorporating an autoregressive moving average (ARMA) model for the errors and allowing the parameters to differ across firms; by estimating via feasible generalized least squares (FGLS), it is possible to obtain consistent estimates for the parameters by allowing returns to correlate over time and across firms. In addition, since we are working with individual stocks, we included the Ibovespa return as a control variable in the model.

In this sense, equation 1 presents the panel autoregressive (PAR) model for stock returns, where c is the constant, $r_{t-j,i}$ are the lagged returns, and $\epsilon_{t,i}$ are the error terms.

$$r_{t,i} = c + \sum_{j=1}^{p}[\beta_j r_{t-j,i} + \theta_j SV_{t-j,i} + \delta_j(r_{t-j,i} \times SV_{t-j-1,i})] + \gamma r_{mt} + \epsilon_{t,i}. \quad (1)$$

Search volume $SV_{t,i}$ is the proxy that represents attention. As mentioned above, Google search volume ($GSV_{t,i}$) and Bloomberg search volume ($BSV_{t,i}$) are used as attention variables. The indices *i*, *t*, and *j* identify each stock, time period, and lag, respectively. We then regress the log returns ($r_{t,i}$) of each stock *i* against its lagged terms ($r_{t-j,i}$), the lagged attention terms ($SV_{t-j,i}$), and the interaction between the lagged return terms and attention terms ($r_{t-j,i} \times SV_{t-j-1,i}$). The order *p* is the maximum lag of the regressors, defined according to Akaike's information criterion, limited to 5 days.

The coefficients $\beta_j$ indicate the relationship between $r_{t,i}$ and its autoregressors. On the other hand, $\delta_j$ indicates how the attention of the previous days affects this relationship. The aim is to test the hypothesis that attention reduces the predictability of returns. Thus, whenever $\beta_j$ shows a significant effect of $r_{t-j,i}$ on $r_{t,i}$, we expect $\delta_j$ to also show a significant but opposite effect of $r_{t,i} \times SV_{t-1,i}$ on $r_{t,i}$. In other words, we expect higher levels of investor attention to reduce the explanatory power of past returns for current returns.

The model allows us to see whether, for example, in the case of corporate events, higher levels of attention are associated with fewer transient shocks due to over- or under-reaction, which would lead to autocorrelation and violate the efficient market hypothesis. Thus, it is a test of the weak form of market efficiency.

Our methodology also includes a specification with trading volume ($TV_{t,i}$) as a control variable. Thus, Model 1 includes returns, search volume, and interactions between the two, while Model 2 also includes trading volume as a fourth regressor. We do not expect trading volume to alter the relationship between attention and return predictability. In addition to trading volume, we include a third specification, adding as control variables the price/earnings (P/E) and market-to-book (MTB) ratios, firm size (measured by the natural logarithm of the market value on each date), stock risk (estimated by the conditional volatility obtained from a standard generalized autoregressive conditional heteroskedasticity (GARCH) model), the Interbank Deposit Certificate (CDI) rate, and the time spread (estimated by the difference between the returns of the 10-year ANBIMA IPCA Constant Duration Index (IDkA) and the 2-year IDkA).

# 4. EMPIRICAL RESULTS

In this section, we describe the sample and present the correlations between the different search volume indicators. We then present the results of the empirical analysis of the effects of investor attention on stock market efficiency.

## 4.1 Descriptive Analysis

Table 2 shows the general characteristics of the variables involved. The final sample consists of 31,960 observations covering 34 stocks and 940 trading days. To avoid the impact of outliers on the analysis, the variables used were winsorized at 1%.

**Table 2**
*Time series descriptive statistics*

| | Return | Volume (in R$ million) | GSV | BSV | Log volume |
|---|---|---|---|---|---|
| **Descriptive statistics for the entire sample** | | | | | |
| Mean | 0.048% | 12.819 | -0.033 | -0 | 15.748 |
| Standard deviation | 2.569% | 16.830 | 0.670 | 1 | 1.250 |
| Median | -0.027% | 7.515 | -0.233 | -0.654 | 15.832 |
| Minimum | -7.590% | 0 | -0.719 | -0.654 | 10.624 |
| Maximum | 8.178% | 490.230 | 3.355 | 2.087 | 18.235 |
| Asymmetry | 0.197% | 4.623 | 2.392 | 1.277 | -1.144 |
| Kurtosis | 1.303% | 46.361 | 7.750 | 0.041 | 3.294 |
| **Descriptive statistics for the average values per share** | | | | | |
| Mean | 0.048% | 12.819 | -0.033 | -0 | 15.748 |
| Standard deviation | 0.091% | 13.022 | 0.264 | 0.712 | 1.069 |
| Median | 0.037% | 8.022 | -0.083 | -0.314 | 15.719 |
| Minimum | -0.280% | 0.773 | -0.375 | -0.654 | 11.989 |
| Maximum | 0.282% | 68.750 | 0.726 | 1.473 | 17.867 |
| Asymmetry | -0.759% | 2.610 | 1.201 | 0.791 | -1.046 |
| Kurtosis | 3.908% | 8.236 | 1.451 | -1.012 | 2.868 |

**Note:** *The main statistical measures of the time series are (i) nominal log returns adjusted for dividends and splits, (ii) monetary volume of transactions (in R$ million), (iii) the index representing the volume of search queries made on Google in Brazil, and (iv) the index representing the volume of readings and search queries on Bloomberg, all for the 34 stocks in the final sample. Transaction volumes are presented in regular and logarithmic form. The top panel shows measures that take into account the entire sample, and the bottom panel shows statistical measures calculated from the average values per share. Daily data from 2018 to 2021 are used.*
*BSV = Bloomberg search volume; GSV = Google search volume.*
**Source:** *Prepared by the authors.*

The measures show an average daily transaction volume of approximately R$12 million. The average daily return is 4.8%, with a SD of 257% and a median of -2.70%. The highest observed daily return was 256.90%, while the lowest was -759%. The highest transaction volume for an individual stock in a day was 490 million, while the highest average daily turnover over the entire period and all stocks was over 68 million. So even after winsorization, there is still a lot of variation in the data. After winsorization, the Google search series no longer have a mean of 0 and a unit SD, indicating high variation in these series.

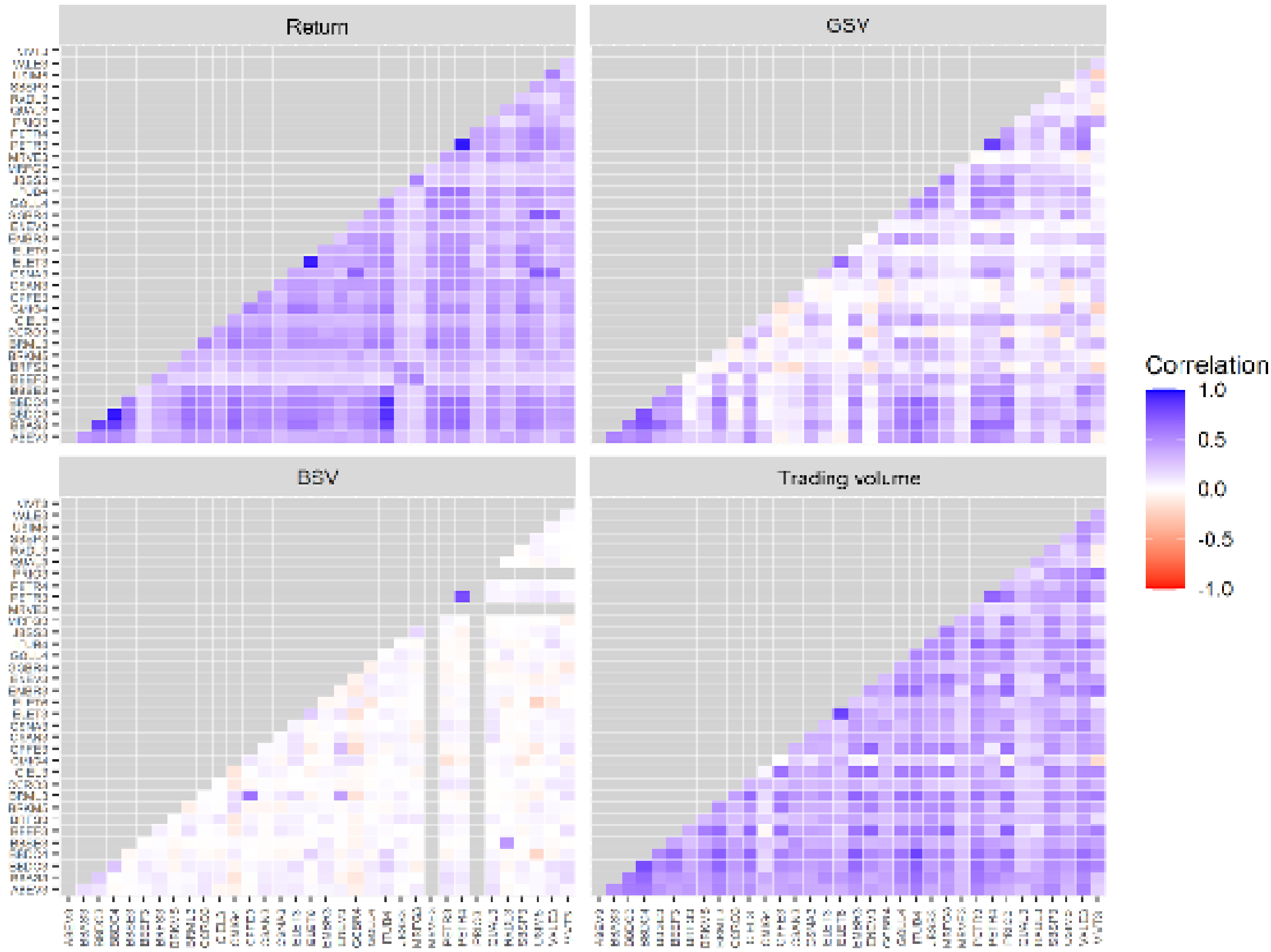


**Figure 1** *Correlation of variables between company pairs*
**Source:** *Prepared by the authors.*

Figure 1 shows the correlation matrices between each pair of stocks for return, transaction volume (in logarithm), and search volume on Google and Bloomberg. Overall, return and volume are correlated between the stocks. On average, returns have a correlation of 43.8%, while the average correlation of volume is 38.4%. Search series, on the other hand, have lower correlations. Google searches have an average correlation of 18.4%, while Bloomberg searches have an average correlation of only 2.4%.

Table 3 shows the correlation between the variables in the analysis, which only captures contemporary correlations, but the temporal dynamics of the variables will be analyzed with the application of the model and presented in the next section. Return has a low correlation with all variables, while trading volume has a relatively high correlation with attention measures (31.17% with Google searches and 28% with Bloomberg searches). Finally, the search indices correlate with each other by only 12.5%, suggesting that the attention levels of professional and retail investors do not always align.

**Table 3**
*Correlation between model variables*

| | Return | log volume | GSV | BSV |
|---|---|---|---|---|
| Return | 1 | | | |
| log volume | 0.0350 | 1 | | |
| GSV | 0.0335 | 0.3171 | 1 | |
| BSV | 0.0274 | 0.2800 | 0.1248 | 1 |

*BSV = Bloomberg search volume; GSV = Google search volume.*
**Source:** *Prepared by the authors.*

In the next section, we present the results of modeling stock returns using autoregressive panel equations. This specification allows us to test the effect of attention on these autocorrelations.

## 4.2 Modeling Results

Table 4 shows the results of the regressions considering non-professional attention, measured by Google searches. The first two models show the estimation of equation 1 including Google searches (a proxy for non-professional attention) in two versions, without and with trading volume lags (in logarithms). The third model shows the results adding the other control variables.

**Table 4**
*Estimation of autoregressive models with Google searches, representing non-professional investor attention*

| | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| | $r_t$ | $r_t$ | $r_t$ |
| (Intercept) | 0.0096 | -0.1001 | -0.2528 |
| | (0.0119) | (0.0987) | (0.2053) |
| $r_{t-1}$ | -0.0155*** | -0.0156*** | -0.0169*** |
| | (0.0047) | (0.0047) | (0.0047) |
| $r_{t-2}$ | -0.0119** | -0.0116** | -0.0116** |
| | (0.0046) | (0.0046) | (0.0047) |
| $r_{t-3}$ | -0.0017 | -0.0017 | -0.0027 |
| | (0.0046) | (0.0047) | (0.0047) |
| $r_{t-4}$ | -0.011** | -0.0115** | -0.0128*** |
| | (0.0046) | (0.0046) | (0.0047) |
| $r_{t-5}$ | 0.006 | 0.0058 | 0.0055 |
| | (0.0047) | (0.0047) | (0.0047) |
| $SV_{t-1}$ | 0.0351* | 0.0349* | 0.0332* |
| | (0.0187) | (0.019) | (0.0192) |
| $SV_{t-2}$ | -0.0226 | -0.0199 | -0.0222 |
| | (0.0219) | (0.0221) | (0.0223) |
| $SV_{t-3}$ | 0.0303 | 0.0298 | 0.0336 |
| | (0.022) | (0.0221) | (0.0223) |
| $SV_{t-4}$ | -0.0153 | -0.019 | -0.0167 |
| | (0.0219) | (0.0221) | (0.0223) |
| $SV_{t-5}$ | -0.0181 | -0.0204 | -0.0185 |
| | (0.0188) | (0.019) | (0.0192) |
| $r_{t-1} \times SV_{t-2}$ | 0.0124** | 0.0124** | 0.0125** |
| | (0.005) | (0.005) | (0.0051) |
| $r_{t-2} \times SV_{t-3}$ | 0.0018 | 0.0019 | 0.0003 |
| | (0.005) | (0.005) | (0.0051) |
| $r_{t-3} \times SV_{t-4}$ | -0.0085* | -0.0083* | -0.0074 |
| | (0.005) | (0.005) | (0.0051) |
| $r_{t-4} \times SV_{t-5}$ | 0.0082 | 0.0083* | 0.0092* |
| | (0.005) | (0.005) | (0.0051) |
| $r_{t-5} \times SV_{t-6}$ | -0.0141*** | -0.014*** | -0.0147*** |
| | (0.0045) | (0.0045) | (0.0045) |
| $TV_{t-1}$ | | -0.0054 | -0.0055 |
| | | (0.0213) | (0.0214) |
| $TV_{t-2}$ | | -0.0235 | -0.0208 |

**Table 4**
*Cont.*

| | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| | $r_t$ | $r_t$ | $r_t$ |
| | | (0.0229) | (0.0231) |
| $TV_{t-3}$ | | 0.0012 | 0.0021 |
| | | (0.023) | (0.0231) |
| $TV_{t-4}$ | | 0.0295 | 0.0295 |
| | | (0.0229) | (0.023) |
| $TV_{t-5}$ | | 0.0051 | 0.0009 |
| | | (0.0212) | (0.0213) |
| $r_m$ | 1.0251*** | 1.0248*** | 0.9976*** |
| | (0.0083) | (0.0083) | (0.0097) |
| *P/E* | | | -0.0002 |
| | | | -0.0002 |
| *MTB* | | | 0.0013 |
| | | | (0.0036) |
| *Size* | | | 0.0086 |
| | | | (0.0104) |
| *Risk* | | | -0.0106 |
| | | | (0.0183) |
| *CDI* | | | 1.7635 |
| | | | (1.6341) |
| *Time spread* | | | 0.156*** |
| | | | (0.0294) |
| Observations | 31,790 | 31,790 | 31,428 |
| Number of shares | 34 | 34 | 34 |
| Number of days | 935 | 935 | 935 |
| Wald statistic | 15,690*** | 15,716*** | 15,911*** |

**Note:** *The models were estimated with a panel via feasible generalized least squares (FGLS) with Parks-Kmenta correction and AR(1) error estimation with Prais-Winsten correction.*
**, **, *** = $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.*
**Source:** *Elaborated by the authors.*

The three models show that there is negative and significant autocorrelation in the daily returns of the stocks analyzed. Thus, the results first indicate that past returns have predictive power for future returns, which contradicts the non-predictability of returns premise of the efficient markets hypothesis (EMH). Second, the estimates of the $\theta_j$ coefficients show that a higher volume of searches on Google is correlated with higher returns the next day, suggesting that attention is greater in optimistic times. The market return (Ibovespa) is highly significant in all specifications. On the other hand, the lags of transaction volume were not significant in any model and, among the other control variables, only the time spread is significant. Finally, the estimates of the $\delta_j$ coefficients, which measure the moderating effect of attention as measured by the volume of searches on Google, show that, in general, greater attention is associated with lower autocorrelation. The first and fifth lags of the interaction are positive and statistically significant, reducing the effect of the negative autocorrelation found.

This result is consistent with our first hypothesis, i.e., that an increase in investor attention promotes greater market efficiency, since higher levels of searches are associated with lower predictability of returns.

Table 5 shows the same specifications, but considering professional attention as measured by Bloomberg searches. The autocorrelation structure of the returns is the same as in the models in Table 4, which also contradicts the HME premise that returns are not predictable, but here there is no relationship between attention and current returns. The control variables also show similar results to those in Table 4, while the interactions between lagged returns and levels of attention, which measure the moderating effect of attention measured by the volume of searches on Bloomberg, are significant (at 10%) at lags 1 and 4, also showing that greater attention is associated with lower autocorrelation, also in line with our first hypothesis.

**Table 5**

*Estimation of autoregressive models with Bloomberg searches representing professional investor attention*

| | Model 4 | Model 5 | Model 6 |
|---|---|---|---|
| | $r_t$ | $r_t$ | $r_t$ |
| (Intercept) | 0.0091 | -0.112 | -0.2585 |
| | (0.0118) | (0.1058) | (0.2111) |
| $r_{t-1}$ | -0.0142*** | -0.0144*** | -0.0156*** |
| | (0.0046) | (0.0046) | (0.0046) |
| $r_{t-2}$ | -0.0108** | -0.0106** | -0.0109** |
| | (0.0046) | (0.0046) | (0.0046) |
| $r_{t-3}$ | -0.0029 | -0.0029 | -0.0038 |
| | (0.0046) | (0.0046) | (0.0046) |
| $r_{t-4}$ | -0.0114** | -0.0118** | -0.013*** |
| | (0.0046) | (0.0046) | (0.0046) |
| $r_{t-5}$ | 0.003 | 0.0029 | 0.0024 |
| | (0.0046) | (0.0046) | (0.0046) |
| $SV_{t-1}$ | 0.0056 | 0.0046 | 0.0046 |
| | (0.0113) | (0.0114) | (0.0115) |
| $SV_{t-2}$ | -0.0085 | -0.0075 | -0.0059 |
| | (0.0124) | (0.0124) | (0.0126) |
| $SV_{t-3}$ | -0.0012 | -0.0017 | 0.0011 |
| | (0.0124) | (0.0125) | (0.0126) |
| $SV_{t-4}$ | 0.0026 | 0.0004 | -0.0024 |
| | (0.0124) | (0.0124) | (0.0126) |
| $SV_{t-5}$ | 0.0037 | 0.0027 | 0.003 |
| | (0.0113) | (0.0114) | (0.0115) |
| $r_{t-1} \times SV_{t-2}$ | 0.0048* | 0.0048* | 0.004 |
| | (0.0029) | (0.0029) | (0.0029) |
| $r_{t-2} \times SV_{t-3}$ | -0.0035 | -0.0036 | -0.0031 |
| | (0.0029) | (0.0029) | (0.0029) |
| $r_{t-3} \times SV_{t-4}$ | -0.0003 | -0.0003 | -0.0006 |
| | (0.0029) | (0.0029) | (0.0029) |
| $r_{t-4} \times SV_{t-5}$ | 0.0055* | 0.0055* | 0.0052* |
| | (0.0029) | (0.0029) | (0.0029) |
| $r_{t-5} \times SV_{t-6}$ | -0.0045 | -0.0045 | -0.0046 |
| | (0.0029) | (0.0029) | (0.0029) |
| $TV_{t-1}$ | | 0.0018 | 0.0007 |
| | | (0.0212) | (0.0214) |
| $TV_{t-2}$ | | -0.0247 | -0.0226 |
| | | (0.023) | (0.0231) |
| $TV_{t-3}$ | | 0.0035 | 0.0043 |
| | | (0.023) | (0.0232) |
| $TV_{t-4}$ | | 0.0281 | 0.0293 |
| | | (0.0229) | (0.023) |
| $TV_{t-5}$ | | -0.0011 | -0.0052 |
| | | (0.0212) | (0.0213) |
| $r_m$ | 1.025*** | 1.0247*** | 0.9972*** |
| | (0.0082) | (0.0082) | (0.0096) |
| *P/E* | | | -0.0002 |
| | | | -0.0002 |
| *MTB* | | | 0.0015 |
| | | | (0.0036) |

**Table 5**
*Cont.*

| | Model 4 | Model 5 | Model 6 |
|---|---|---|---|
| *Size* | | | 0.0085 |
| | | | (0.0104) |
| *Risk* | | | -0.0091 |
| | | | (0.0182) |
| *CDI* | | | 1.6074 |
| | | | -1.5906 |
| *Time spread* | | | 0.1572*** |
| | | | (0.0292) |
| Observations | 31,756 | 31,756 | 31,394 |
| Number of shares | 34 | 34 | 34 |
| Number of days | 934 | 934 | 394 |
| Wald statistic | 15,819*** | 15,864*** | 16,063*** |

**Note:** *The models were estimated with a panel via feasible generalized least squares (FGLS) with Parks-Kmenta correction and AR(1) error estimation with Prais-Winsten correction.*
**, **, *** = p < 0.1, p < 0.05, p < 0.01, respectively.*
**Source:** *Prepared by the authors.*

In both Table 4 and Table 5, the interactions between lagged returns and attention have the opposite sign to the autocorrelation coefficients, indicating that the level of attention reduces the predictability of returns, consistent with the first hypothesis of the paper that attention contributes to greater efficiency. To better understand the effect of different levels of attention on the predictability of returns, we calculated the partial effect of lagged returns $\left(\partial r_t / \partial r_{t-j} = \beta r_{t-j} + \delta\left(r_{t-j} \times \mathrm{SV}_{t-j-1}\right)\right)$ for different values of the SV variable and for the lags whose interaction was significant in the models in tables 4 and 5. The results of this analysis are presented in Table 6.

**Table 6**
*Autocorrelation (partial effects) for different levels of attention*

| Non-professional attention (Google searches) | | | | |
|---|---|---|---|---|
| **Lag** | **Level of attention** | **Model 1** | **Model 2** | **Model 3** |
| $r_{t-1}$ | Low attention | -0.0212 | -0.0212 | -0.0226 |
| $r_{t-1}$ | Median attention | -0.0184 | -0.0185 | -0.0198 |
| $r_{t-1}$ | High attention | -0.0133 | -0.0134 | -0.0147 |
| $r_{t-5}$ | Low attention | 0.0064 | 0.0064 | 0.0067 |
| $r_{t-5}$ | Median attention | 0.0033 | 0.0033 | 0.0034 |
| $r_{t-5}$ | High attention | -0.0025 | -0.0025 | -0.0026 |
| **Professional attention (Bloomberg searches)** | | | | |
| **Lag** | **Level of attention** | **Model 4** | **Model 5** | **Model 6** |
| $r_{t-1}$ | Low attention | -0.0173 | -0.0175 | -0.0156 |
| $r_{t-1}$ | Median attention | -0.0173 | -0.0175 | -0.0156 |
| $r_{t-1}$ | High attention | -0.0140 | -0.0142 | -0.0156 |
| $r_{t-4}$ | Low attention | -0.0139 | -0.0143 | -0.0153 |
| $r_{t-4}$ | Median attention | -0.0127 | -0.0131 | -0.0142 |
| $r_{t-4}$ | High attention | -0.0105 | -0.0109 | -0.0120 |

**Note:** *Low attention is the value of the first quartile of the attention variable (-0.4544 for Google searches and -0.6536 for Bloomberg searches), median attention is the value of the median (-0.2328 for Google searches and -0.6536 for Bloomberg searches), and high attention is the value of the third quartile (0.1798 for Google searches and 0.0316 for Bloomberg searches).*
**Source:** *Prepared by the authors.*

In the first part of Table 6, we show the effect of different levels of non-professional attention (Google searches) on the order 1 and 5 autocorrelations. In the three specifications (Models 1 to 3 in Table 4), the autocorrelation decreases as attention increases. For example, in Model 3, the order 1 autocorrelation with low attention is -0.0226. When attention is high, the autocorrelation decreases to -0.0147. In the second part of Table 6, we show the effect of different levels of professional attention (searches on Bloomberg) on the order 1 and 4 autocorrelations. In order

1, only models 4 and 5 show a variation in predictability as a function of attention, which also indicates a reduction in predictability in these cases. In order 4, all specifications show a decrease in predictability with increasing attention. For example, in Model 6, the autocorrelation is -0.0153 when attention is low and decreases to -0.0120 when attention is high. However, the results suggest that the effect of professional attention is less than that of non-professional attention, since the decrease in predictability is less in models 4, 5, and 6 than in models 1, 2, and 3, contrary to the second hypothesis of this study.

A possible explanation for these results is the low variability of the Bloomberg searches variable, which measures professional attention. As can be seen in Table 2, the median and the minimum have the same value for the BSV index. The GSV index (non-professional attention) shows a higher variability (there are fewer observations concentrated in the same values). Another plausible explanation is that, due to their behavioral biases, non-professional investors play a more decisive role in market transients than professional investors. However, these transients would be more present (or more expressive) in moments of inattention rather than attention on the part of non-professional investors.

### 4.2.1 Additional analyses

Since the time spread was the only significant control variable for returns in tables 4 and 5, in this section we analyze whether different levels of spread, which measures uncertainty about the future at the macroeconomic level (the higher the spread, the greater the uncertainty), affect the relationship between attention and the predictability of returns. For this test, we estimated a regression including triple interactions between lagged returns, attention, and the time spread. While the time spread is a measure of macroeconomic risk, we also considered a measure of the individual risk of each stock, the *Risk* variable used as a control in models 1 to 6. The results of these estimations are not presented due to space constraints; we only show the results of the partial effects of lagged returns (autocorrelation) for different levels of attention and spread and individual risk in Table 7. In the first part of Table 7, we show the partial effects of Google searches for the first lag of returns, because this was the lag for which the three-way interaction was significant for both spread and individual risk. As for Bloomberg searches, the three-way interaction for spread and individual risk was significant for the fourth lag, so this is the lag analyzed in the second part of Table 7.

**Table 7**
*Autocorrelation (partial effects) for different levels of attention, spread, and individual risk*

| | | **Non-professional attention (Google searches)** | | |
|---|---|---|---|---|
| **Lag** | **Level of attention** | **Low spread** | **Median spread** | **High spread** |
| $r_{t-1}$ | Low attention | -0.0208 | -0.0259 | -0.0310 |
| $r_{t-1}$ | Median attention | -0.0202 | -0.0228 | -0.0254 |
| $r_{t-1}$ | High attention | -0.0190 | -0.0170 | -0.0149 |
| | | **Low risk** | **Median risk** | **High risk** |
| $r_{t-1}$ | Low attention | -0.0252 | -0.0226 | -0.0187 |
| $r_{t-1}$ | Median attention | -0.0195 | -0.0181 | -0.0161 |
| $r_{t-1}$ | High attention | -0.0088 | -0.0098 | -0.0114 |
| | | **Professional attention (Bloomberg searches)** | | |
| **Lag** | **Level of attention** | **Low spread** | **Median spread** | **High spread** |
| $r_{t-4}$ | Low attention | -0.0129 | -0.0147 | -0.0165 |
| $r_{t-4}$ | Median attention | -0.0129 | -0.0147 | -0.0165 |
| $r_{t-4}$ | High attention | -0.0110 | -0.0109 | -0.0108 |
| | | **Low risk** | **Median risk** | **High risk** |
| $r_{t-4}$ | Low attention | -0.0127 | -0.0144 | -0.0170 |
| $r_{t-4}$ | Median attention | -0.0127 | -0.0144 | -0.0170 |
| $r_{t-4}$ | High attention | -0.0134 | -0.0134 | -0.0132 |

**Note:** *The levels of attention are the same as those defined in Table 6. Low spread is measured by the first quartile of the spread variable (-0.2553), median spread is a zero value, and high spread is the third quartile (0.2590). Low risk is measured by the first quartile of the risk variable (1.9200), median risk is the median of the variable (2.8200), and high risk is the value of the third quartile (2.8160).*
**Source:** *Prepared by the authors.*

In the first part of Table 7 (Google searches, i.e. non-professional attention), it is possible to see a greater autocorrelation as the spread increases, suggesting that uncertainty about the future hampers the informational efficiency of the market. For a low spread, greater attention generates slightly lower autocorrelations, but when the spread is high, the autocorrelation drops by half as attention goes from a low level (-0.0310) to a high level (-0.0149). For individual risk, the interpretation is similar: when attention is low or at the median level, the greater the risk, the greater the autocorrelation, and for high levels of risk, the effect of attention is stronger (the autocorrelation with low attention is -0.0187 and with high attention it is -0.0114).

In the second part of Table 7 (searches on Bloomberg, i.e. professional attention), as the spread increases (when attention is low or at the median level), it is possible to see greater autocorrelation, suggesting that uncertainty about the future hampers the informational efficiency of the market. When the spread is low, greater attention generates slightly lower autocorrelations, but when the spread is high, the autocorrelation decreases more as attention goes from a low level (-0.0165) to a high level (-0.0108). For individual risk, the interpretation is similar for high levels of risk: when attention is low or at the median level, the autocorrelation is -0.0170, and when attention is high, the autocorrelation drops to -0.0132. Comparing the decrease in autocorrelation for professional attention and non-professional attention, the conclusion regarding the second hypothesis of the study stands: the effect of non-professional attention is greater, probably due to the limitations of the professional attention measure as commented in the previous section.

### 4.2.2 Robustness analyses

Finally, these results are robust to several different specifications. We estimated the regressions with data without winsorization and winsorized at 5%, with the logarithm in the search volume variable, and also using shorter periods (3 and 2 years). In addition, in our main specification, attention is one lag order below returns for the analysis of autocorrelations (Tantaopas et al., 2016), but we also re-estimated all models including attention at the same lag order as returns (Vozlyublennaia, 2014) and the conclusions remained the same. The results of these analyses are not presented here due to space constraints, but are available upon request.

## 5. CONCLUDING REMARKS

In this study, we investigated whether market efficiency is affected by the level of investor attention. Efficiency is measured by a lower predictability of returns resulting from a more immediate incorporation of information into prices. The level of internet activity is used as a measure of attention. Searching the internet is currently a recurring process for keeping informed about assets traded on the stock exchange. Since it is not directly linked to the market, this measure is not the result of equilibrium, which minimizes endogeneity problems.

We used measures of searches made on both Google and Bloomberg. While Google is popular, free, and aimed at non-professional investors, Bloomberg is aimed at more sophisticated investors. This allows us to test whether the effects of attention depend on which agent is paying attention. According to information discovery theory (Tantaopas et al., 2016; Vozlyublennaia, 2014), greater investor attention should lead to greater efficiency, and our results are consistent with this hypothesis. However, contrary to our complementary hypothesis, we find no evidence that professional attention plays a more decisive role than non-professional attention in inducing efficiency. Contrary to what we originally hypothesized, the evidence suggests that phenomena such as bounded rationality and herd behavior, which have been documented in previous studies (Chiang & Zheng, 2010; Kahneman, 2003; Sewell, 2007), may be more present in times of low than in times of high non-professional investor attention. In addition, although these investors are less informed on average, they are more susceptible to adverse selection due to information asymmetry (Akerlof, 1970; Cohen et al., 2011).

To conduct the analysis, we obtained daily data from 2018 to 2021 for 34 individual stocks. This resulted in 935 trading days, which is sufficient from a statistical point of view. As far as we have seen, our work is pioneering in carrying out this type of test on a daily basis. We studied some of the main Brazilian stocks, starting with the Ibovespa. Also for the first time, such an analysis of the incorporation of information into the prices of risky assets is carried out by separating professional and non-professional attention. We chose to study individual stocks rather than indices because there is less literature on this universe. In addition, the search volume index presented by Bloomberg is only available for stocks.

For this analysis, we developed autoregressive models of returns that included as regressors interactions between these lagged returns and internet search indices for the stock ticker. A minimizing effect on the predictability of returns was found, considering both professional and non-professional attention. The information discovery process is the mechanism by which attention induces efficiency in the markets. This process begins with a significant price variation resulting from a relevant event. This fact attracts the attention of some investors, who then search for more information to support their decision to buy, sell, or hold the asset. This is a process that naturally transmits information to prices, and attention is a prerequisite for it to occur.

We aimed to make some contributions by developing a pioneering analysis of the effect of the attention of different classes of investors on market predictability. First, we acknowledge the importance of attention and action by professional investors in incorporating information into prices. In addition to contributing to the literature on the market efficiency hypothesis, the analysis involving the volume of internet searches is particularly relevant given the large amount of information that people now produce and share digitally. The results show that the attention of investors with different profiles has different effects on price predictability.

Our study has some limitations that could be addressed in future research. From a methodological point of view, we selected only companies that are part of the Ibovespa, only attention on trading days, and eliminated extreme values from the attention series. With regard to the theoretical argument, variations in risk (and consequently in risk premiums) limit the relationship between predictability of returns and market efficiency. We tried to circumvent this limitation by controlling for volatility, but we encourage the search for alternative ways to establish this relationship.

Future research could also look at the implementation of empirical models to better understand the mechanisms by which attention induces market efficiency. Another promising avenue to complement our findings is to evaluate the possibility of gains from investment strategies that take into account the greater predictability of stock prices during periods of reduced attention.

## REFERENCES


Akerlof, G. A. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *The Quarterly Journa of Economics*, *84*(3), 488-500. https://doi.org/10.2307/1879431

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31-56. https://doi.org/10.1016/S1386-4181(01)00024-6

Andrei, D., & Hasler, M. (2015). Investor attention and stock market volatility. *Review of Financial Studies*, *28*(1), 33-72. https://doi.org/10.1093/rfs/hhu059

B3 S.A. – Brasil, Bolsa, Balcão. (2015). *Metodologia do índice Bovespa*. https://www.b3.com.br/data/files/1C/56/F7/D5/96E615107623A41592D828A8/IBOV-Metodologia-pt-br.pdf

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional in. *Review of Financial Studies*, *21*(2), 785-818. https://doi.org/10.1093/rfs/hhm079

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, *49*(3), 307-343. https://doi.org/10.1016/S0304-405X(98)00027-0

Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies*, *30*(9), 3009-3047. https://doi.org/10.1093/rfs/hhx031

Burton, E. T., & Shah, S. N. (2013). *Behavioral finance: Understanding the social, cognitive, and economic debates* (Vol. 854). John Wiley & Sons.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. Cambridge University Press.

Castro, F. H., Edi Jr., W., Santana, V. F., & Yoshinaga, C. E. (2019). Fifty-year history of the Ibovespa. *Revista Brasileira de Finanças*, *17*(3), 47-65.

Chiang, T. C., & Zheng, D. (2010). An empirical analysis of herd behavior in global stock markets. *Journal of Banking & Finance*, *34*(8), 1911-1921. https://doi.org/10.1016/j.jbankfin.2009.12.014

Cohen, J., Holder-Webb, L., Nath, L., & Wood, D. (2011). Retail investors’ perceptions of the decision-usefulness of economic performance, governance, and corporate social responsibility disclosures. *Behavioral Research in Accounting*, *23*(1), 109-129. https://doi.org/10.2308/bria.2011.23.1.109

Cronqvist, H., & Siegel, S. (2014). The genetics of investment biases. *Journal of Financial Economics*, *113*(2), 215-234. https://doi.org/10.1016/j.jfineco.2014.04.004

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461-1499. https://doi.org/10.1111/j.1540-6261.2011.01679.x

Falkinger, J. (2008). Limited attention as a scarce resource in information-rich economies. *The Economic Journal*, *118*(532), 1596-1620. https://doi.org/10.1111/j.1468-0297.2008.02182.x

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, *38*(1), 34-105.

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance*, *64*(5), 2023-2052. https://doi.org/10.1111/j.1540-6261.2009.01493.x

Gabaix, X., & Laibson, D. (2005). *Bounded rationality and directed cognition*. Harvard University.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, *70*(3), 393-408.

Grullon, G., Kanatas, G., & Weston, J. P. (2004). Advertising, breadth of ownership, and liquidity. *The Review of Financial Studies*, *17*(2), 439-461. https://doi.org/10.1093/rfs/hhg039

Johansson, E. (2014). *Creating daily search volume data from weekly and daily data.* https://erikjohansson.blogspot.com/2014/12/creating-daily-search-volume-data-from.html

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, *93*(5), 1449-1475. https://doi.org/10.1257/000282803322655392

Lux, T., & Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, *397*(6719), 498-500.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, *42*(3), 483-510.

Mondria, J., & Wu, T. (2010). The puzzling evolution of the home bias, information processing and financial openness. *Journal of Economic Dynamics and Control*, *34*(5), 875-896. https://doi.org/10.1016/j.jedc.2009.12.004

Sewell, M. (2007). *Behavioural finance.* https://www.academia.edu/2813323/Behavioural_Finance

Sims, C. A. (2005). *Rational inattention: A research agenda.* https://www.bundesbank.de/resource/blob/703284/82d389a1de62876b9151575e466767cd/mL/2005-09-27-dkp-34-data.pdf

Statista. (2014). *Popular online search engines in Brazil as of May 2014, based on market share.* https://www.statista.com/statistics/309652/brazil-market-share-search-engine/

Tantaopas, P., Padungsaksawasdi, C., & Treepongkaruna, S. (2016). Attention effect via internet search intensity in Asia-Pacific stock markets. *Pacific-Basin Finance Journal*, *38*, 107-124. https://doi.org/10.1016/j.pacfin.2016.03.008

Van Nieuwerburgh, S., & Veldkamp, L. (2007). *Information immobility and the home bias puzzle* [Working Paper]. Wiley Online Library. https://doi.org/10.1111/j.1540-6261.2009.01462.x

Vozlyublennaia, N. (2014). Investor attention, index performance, and return predictability. *Journal of Banking & Finance*, *41*, 17-35. https://doi.org/10.1016/j.jbankfin.2013.12.010

Zhang, W., Shen, D., Zhang, Y., & Xiong, X. (2013). Open source information, investor attention, and asset pricing. *Economic Modelling*, *33*, 613-619. https://doi.org/10.1016/j.econmod.2013.03.018