Investor attention and Portuguese stock market volatility: We'll google it for you!

ANA BROCHADO

Instituto Universitario de Lisboa (ISCTE - IUL), Centre for Socioeconomic and Territorial Studies (DINÂMIA'CET – IUL)

1. INTRODUCTION

The importance of investor attention in financial markets is well established on a theoretical level (e.g. Merton, 1987; Hirshleifer and Teoh, 2003; Sims, 2003), and several proxies for investor attention have already been proposed (e.g. Barber and Odean, 2008). More recently, Da et al. (2014) conducted one of the first studies to incorporate Internet search behaviour as a proxy for retail investors' attention allocation. This approach recognises that the Internet has become a mainstream platform for the production, intermediation and consumption of information in the financial industry. Search engines are an intuitive research tool that provides access to huge amounts of information at a negligible cost. Investors consider information attention a valuable cognitive resource (Zhang et al., 2013), and investors who pay attention to stock or market indices habitually search for new information about them. Weng et al. (2018) maintain that Google search data, in particular, capture traders' collective interest.

Studying investor attention is thus of utmost importance since previous studies (e.g. Aouadi et al., 2013) have concluded that it acts as a significant determinant of stock activity. More specifically, Mondria and Quintanna-Domeque (2012) argue that investor attention is a new transmission channel of financial crises across markets.

The present study sought to contribute to the literature by providing further evidence that investor attention is a determinant of stock market volatility. In addition, for the first time, Portuguese Google volume search data were analysed. Accordingly, this study's objectives were twofold. First, this research investigated the effect on stock market volatility of information attention at both the individual stock and overall market levels. Second, the study attempted to test whether this relationship remains stable across different market states.

This paper's remaining sections are as follows. The next section discusses the concepts of the Internet and investor attention and provides a literature review of previous studies that have used Internet search data in financial contexts. Then, the methodological options selected are presented, namely, Google search data and stock market activity variables. The results section reports the main model estimates and discusses the findings. We finish with conclusions and proposed directions for future research.

2. LITERATURE REVIEW

2.1 *Internet and Investor Attention*

Merton (1987) introduced the concept of investor recognition, suggesting that investor attention might be an important determinant of stock market activity. Attention is known to play an important role in investors' learning and trading behaviour. However, the exact role of information and investor attention in market efficiency remains elusive. Grossman and Stiglitz's (1980) seminal work maintains that more information leads to more informative prices, which should improve market efficiency. On the other hand, Da et al. (2010) argue that more attention can create extra noise and reduce market efficiency.

Regardless, the assumption that investor attention influences stock market activity is supported empirically by a number of studies that propose different proxies for attention. For example, Barber and Odean (2008) describe 'attention-grabbing stocks' as those stocks that first capture investors' notice. News headlines, high abnormal trading volume and extreme returns can retain investors' attention. Consequently, stocks receiving more attention from investors become relatively more traded than do those attracting less attention. As investors appear to limit their attention to a small number of stocks (i.e. limited attention bias), a delay occurs in investor response, and new information is not instantaneously incorporated into stock prices (Mondria et al., 2010). Investor attention seems to interact with cognitive bias affecting investors' reactions to new information. However, even though financial information acquired by investors is not fully incorporated into stock prices, Internet searches can enhance the speed of information dissemination, thereby making the market more efficient (Zhang et al., 2013).

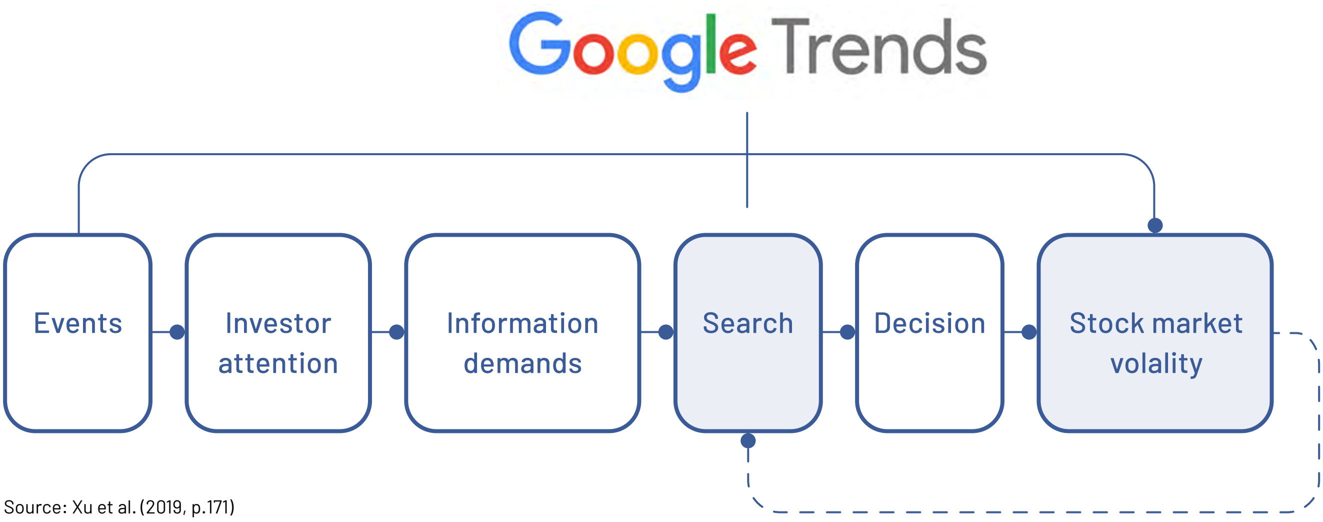
Previously developed proxies for investor attention include companies' brand perception (Frieder and Subrahmanyam, 2005), advertising expenditure (Grullon et al., 2004; Chemmanur and Yan, 2009; DellaVigna and Pollet, 2009), media coverage (Barber and Odean, 2008; Fang and Peress, 2009; Yuan, 2011), trading volume (Barber and Odeon, 2008) and price limit events (Seasholes and Wu, 2007). However, measuring investor recognition is still a difficult task, and some shortcomings are associated with the above-listed measures (Bank et al., 2011). For instance, no reliable information exists about the extent to which newspaper readers pay attention to mentions of companies in articles or consumers pay attention to companies' advertising activity. Furthermore, excessive trading volume and stock returns are also determined by macroeconomic variables, which are unrelated to investor attention (Zhang et al., 2013).

Da et al.'s (2014) innovative study tested the use of Google search volume for ticker symbols as a proxy for firm-specific investor attention. Google search volume data are made available by Google Trends, an online tool that provides access to the relative online search volume for any query term submitted to Google by Internet users. The cited authors' results confirm that search-based data act as a more direct and timely proxy for attention than previously used proxies do, including extreme returns, trading volume and media attention. Moreover, the results reveal that Google search volume is positively associated with market capitalisation, abnormal returns, turnover and media attention.

Bank et al. (2011) point out some relative advantages of using Google search volume as a direct measure of investor attention. First, information is considered a valuable resource in financial markets. As the World Wide Web is the largest pool of freely available information, a not unrealistic conclusion is that most investors are regular users of the Internet. Google Trends' data on searches is a particularly useful data source due to the popularity of Google as a search engine. Second, search volume data originate with both investors and customers. An Internet user clearly only makes a specific query if he or she is interested in the object underlying the search term. Third, search volume data are reported and updated on a daily basis. Therefore, these data capture not only individual investors' attention but also a timelier version of this than other well-established attention variables do. Thus, Internet search queries can be interpreted as a measure of retail investors' attention in the stock market, as suggested by Da et al. (2014).

According to Xu et al. (2019), the search volume for companies is associated with market events' impact. Events are expected to attract investors' attention in an extremely short time. Investors subsequently search for information on the Web before deciding how to respond to each event. Events can thus cause stock market volatility, providing a rational basis for believing that a relationship exists between Internet searches and stock market volatility.

Figure 1: Events and Google Trends' mechanisms of influence on stock volatility



Source: Xu et al. (2019, p.171)

2.2 *Overview of Empirical Studies*

A recent but growing stream of literature has highlighted the power of Google search volume in a variety of settings in financial research. In particular, one strand of research has suggested that Google search volume helps to explain and predict stock market activity.

Mondria et al. (2010) conducted one of the first studies that incorporated the behaviour of Internet search engine users as an indicator of investors' attention allocation. The cited authors combined United States (US) data on portfolio holdings of foreign securities with the attention of American investors in search queries focused on these foreign markets. The results reveal that investors look for more information about countries where these individuals hold more assets and that investors are more active in countries about which these individuals are more informed.

Da et al. (2014) did another pioneering study proposing the use of search volume data in financial applications. The cited authors found that Internet searches for firms' most popular products are positively correlated with revenue surprises and that Internet search data are an interesting option as a proxy for investor attention.

Da et al. (2011) also provide evidence that Internet search data for assets' ticker symbols capture retail investor attention in a timelier and more accurate way than other proxies for investor attention. The cited study's findings further reveal that an increase in Internet searches for Russell 3000 stocks predicts higher stock prices in the next two weeks and a possible price reversal within the year. Moreover, Internet search data are also associated with large first-day returns and long-run underperformance for the initial public offering stocks under study.

Bank et al. (2011) suggest that Google search volume acts as an intuitive proxy for firm recognition and accurately portrays stock market investors' attention. The cited authors report that an increase in Google search volume for a company's name is associated with a rise in trading activity and stock liquidity related to the same firm – at least in the short run. Bank et al. (2011) posit that the observed positive relationship between search volume and liquidity is most likely due to changes in the cost of asymmetric information. The latter authors argue that search volume primarily measures uninformed investors' interest. In addition, Bank et al. (2011) offer evidence that an increase in attention is associated with short-term buying pressure and that this leads to temporarily higher returns.

Dimpfl and Jank (2011) investigated the dynamics of stock market volatility and retail investor attention to the aggregate stock market. The data for the cited study encompassed the realised volatility of four leading market indexes – the Dow Jones, FTSE, CAC and DAX. Investor attention was measured by the search activity focused on their respective names. The results indicate the existence of a bi-directional Granger causality between the realised volatility of stock market indices and Internet search queries. Thus, investor attention to the stock market grows in times of high market movements, and a rise in investor attention is accompanied by higher volatility. Search query data, thus, have predictive power for future stock market volatility.

Dzielinski (2012) reports the existence of a positive relationship between the frequency of Internet searches for the word 'economy' and aggregate stock returns and volatility. The cited author also asserts that this proposed measure is correlated with other measures of confidence and uncertainty. Therefore, Dzielinski (2012) maintains that the volume of Internet searches with the word 'economy' as their topic needs to be considered a measure of uncertainty about the state of the economy rather than a measure of investor attention. As economic psychology posits that a higher degree of uncertainty about the economy increases the demand for information, the cited author argues that agents respond to increased uncertainty by intensifying their volume of Internet searches with 'economy' as their topic.

Drake et al. (2012) investigated investor search behaviour centred around corporate announcements. The cited authors also studied how variations in Internet searches impact the capital market's response to earnings. According to Drake et al.'s (2012) results, the volume of Internet searches rises to abnormal levels about two weeks prior to earnings announcements, spikes markedly at the announcement and stays at high levels for some time after announcements. Internet searches are positively related to news and media attention and are negatively associated with investor distraction. Internet searches, thus, can partially anticipate the information content of earnings news.

Vlastakis and Markellos (2012) report evidence for an economically positive and robust link between Internet search data at the market level and implied volatility, historical volatility and trading volume, even after controlling for market return and information supply. The cited authors also examined the stability of these results across different market scenarios, concluding that Internet search data increase significantly during periods of higher returns. Finally, these authors observe that investors use more active Internet search data as their level of risk aversion increases.

A more recent study by Aouadi et al. (2013) delved into investor attention's influence on French stock market activity, liquidity and volatility. The results indicate that Google search volume for firms' names is strongly correlated with trading volume and stock liquidity. In addition, the identified links remain economically stable over time, even after controlling for the effects of financial crises. Aouadi et al. (2013) emphasise that using the Internet to acquire financial information accelerates information dissemination about stock prices and helps to increase the stock market's efficiency.

Zhang et al. (2013) conclude that investor attention is a statistically significant explanatory variable for abnormal returns, even after controlling for trading volume. The cited authors focused on three different electronic boards of the Chinese market and confirmed the existence of a bi-directional Granger causality between Internet search data and abnormal returns.

Latoeiro and Ramos (2013) studied whether Internet search queries can predict stock market activity. These authors argue that investors incorporate more market information than stock specific information in their investment decisions. The results hold true for both the market index and stock levels. An increase in net searches for stocks leads to an increase in volatility and volume and a decrease in cumulative returns. Moreover, an increase in Internet searches for market indexes is followed by a decrease in the indexes' returns and an increase in implied volatility.

Chouliaras and Grammatikos (2013) report a positive effect of the web attention index on a country's economy, resulting in the probability of extreme returns for different European countries in particular. In addition, more web attention in times of crises is associated with a higher probability of extreme bottom returns.

Takeda and Wakao (2014) suggest that a stronger correlation exists between search intensity and trading volume than between search intensity and stock returns. The correlation between search intensity and returns also appears to be stronger for smaller stocks. The effect of search intensity on trading volume is not affected by investor sentiment.

Vozlyublennaia (2014) presents an analysis of the link between investor attention and the performance of several indexes in broad investment categories. The empirical results reveal a short-term change in index returns following an increase in attention and a long-term change in attention as a result of a shock to returns. According to the cited author, investor attention diminishes index returns' predictability and improves market efficiency.

Heiberger (2015) examined the Google search volumes for all companies listed on the Standard & Poor's 100 Index comprising the major branches of the Global Industry Classification Standard taxonomy. The cited author observes that collective behaviour indicators can predict recessions since downturns draw mass attention to the economic outlook.

Lobão et al. (2017) studied firms' profitability against market trends for Europe's top-of-mind stocks. The cited researchers suggest only weak evidence is available that an increase in firms' Google search volumes can result in an abnormal return. Investors thus gain a better return if they adopt a passive investment strategy.

Weng et al. (2018) present a financial expert system to predict short-term stock prices that combines different data sources including Google Trends, financial news, stock market data, Wikipedia data and technical indicators employing different machine learning methods. The cited authors report that search terms and other online sources can be used to supplement traditional financial data to improve their expert system's performance. Xu et al. (2019), in turn, argue that mixed-frequency data (i.e. low frequency data on economic fundamentals and high event data based on Google search data) can be used to predict monthly volatility in the US stock market.

Table 1 gives an overview of the main methodological options of the reviewed studies, including sample characteristics, dependent variable in the analysis, Internet search data options and other control variables. Whereas some studies have analysed aggregate stock market activity (e.g. Dimpfl and Jank, 2011;

Dzielinski, 2012; Vozlyublennaia, 2014), other studies have assessed investor attention's influence at the individual stock level (e.g. Da et al., 2011; Drake et al., 2012; Vlastakis and Markellos, 2012; Aouadi et al., 2013; Weng et al., 2018). As regards a proxy for investor attention, previous studies have used two methodological options. For example, Da et al. (2011) and Drake et al. (2012) defend the use of an Internet search volume based on stock ticker symbols. Other authors (e.g. Bank et al., 2011; Vlastakis and Markellos, 2012; Lobão et al., 2017) prefer using Internet search data based on company names.

Table 1: An overview of financial studies that use Internet search data as a proxy for investor attention

Reference	Sample	Stock Market Variables	Internet Search Data	Other Variables & Control Variables	Statistical Methods
Mondria et al. (2010)	The US, March 2006	Attention, holdings (country level)	Number of times users searched results for a particular country	Market capitalisation, English-speaking dummy, female dummy, per capita gross domestic product, bilateral trade with the U.S., distance to the U.S.	Regression analysis, two-stage least squares, three-stage least squares
Da et al. (2014)	865 firms listed by Nielsen Media Research (i.e. those that advertised a product on television), 2004-2008.	Revenue surprise, announcement returns, earnings informativeness, earnings management, post-announcement returns	SVI of firms' most popular products	Size, market-to-book, turnover, prior return, institutional ownership	Panel data
Da et al. (2011)	Russell stocks, 2004-2008	SVI and abnormal SVI	SVI for stock ticker symbols and company name, abnormal SVI	Stock and abnormal returns, turnover and abnormal turnover, market capitalisation, news-based data, advertising expenses/sales	VAR models, panel data
Dimpfl and Jank (2011)	Stock market indexes: Dow Jones, FTSE, CAC and DAX, July 2006-June 2011	Aggregate stock market volatility and google search data	SVI on stock market indexes' names	Aggregate stock market volatility and Google search data (lagged)	Granger causality test, VAR models

Reference	Sample	Stock Market Variables	Internet Search Data	Other Variables & Control Variables	Statistical Methods
Bank et al. (2011)	German stock market, Xetra trading system, January 2004–June 2010	Liquidity, turnover	SVI on firm's names	Return, turnover, interaction between stock traded volume and Savills share price, liquidity (lagged variables)	Stock portfolio formation, panel data
Dzielinski (2012)	Market indexes from the US, Australia, Canada, the United Kingdom, Germany and Japan, January 2005–June 2011	Aggregate stock returns S&P 500 Index, weekly realised volatility	SVI for 'economy'	Indicators of confidence and uncertainty, economic crisis (May 2007–June 2009)	Correlation analysis, regression analysis
Drake et al. (2012)	Stocks of the S&P 500, 2005–2008	Abnormal search volume	SVI on tickers' names, daily data	Earnings announcements, management, forecast date, analysis forecast data, dividend announcement, acquisition announcement, return, turnover, bid-ask spread, institutional ownership, firm attributes	Correlation analysis, regression analysis
Vlastakis and Markellos (2012)	30 of the largest stocks traded on the NYSE and NASDAQ	Realised volatility, implied volatility, expected variance risk premium	SVI on companies' names	Market return, firm-specific information supply (firm level, market level)	Correlation analysis, hypothesis testing, GARCH models, regression, panel data
Zhang et al. (2013)	Shanghai Stock Exchange and Shenzhen Stock Exchange, 30 stocks from Chinext, 30 from the SME exchange, 30 from the Main Board, March 2011–March 2012	Abnormal returns	BI on stocks' names, daily data	Trading volume	Granger causality tests, regression analysis
Aouadi et al. (2013)	Stocks from the CAC, France, January 2004–June 2010	Illiquidity, volatility	SVI on firms' names	Volatility, return, stock traded volume, trend, interaction between stock traded volume and SVI, global crisis effect	Correlation analysis, unit root tests, regression analysis, panel data

Reference	Sample	Stock Market Variables	Internet Search Data	Other Variables & Control Variables	Statistical Methods
Latoeiro and Ramos (2013)	Stocks from the Euro Stoxx 50 Index, January 2004–June 2011	Volume, abnormal volume, returns, absolute returns, cumulative returns, historical and implied volatility	SVI on firms and indexes' names	Lagged returns, volatility, volume	Regression, panel data
Chouliaras and Grammatikos (2013)	Stock market indexes from three groups of European countries: (i) euro-periphery countries (ii) euro-core countries and (iii) major European countries – but not euro-countries, January 2004–March 2013	Returns	Web SVI for [country] crisis, [country] debt, [country] economy, [country] deficit, [country] default	Pessimist news factor, news relevance factor, stock returns	Granger causality test, VAR models
Takeda and Wakao (2014)	Japanese stocks	Returns, volume, abnormal returns, abnormal volume	SVI on company names	Interaction term between search intensity and investor sentiment (volatility index)	Portfolio analysis, panel data
Vozlyublennaia (2014)	Security indexes in broad investment categories, Dow Jones Industrial Average (Dow), NASDAQ, S&P500 10-year Treasury index, Chicago Board Options Exchange Gold index, West Texas Intermediate crude oil index, January 2004–December 2012	Returns, volatility	SVI on indexes' names	Interaction terms between lagged attention and lagged returns, macroeconomic variables, default spread, one-year Treasury bill rate aggregate dividend yield	Granger causality test, VAR models, regression analysis
Heiberger (2015)	S&P100, January 2004–January 2014	Returns (by sector)	SVI for company names	Volatility index	Regression analysis
Lobão et al. (2017)	Stocks from the Euro Stoxx, 50 Index	Portfolio returns	SVI for company names	n/a	Portfolio and regression analyses

Reference	Sample	Stock Market Variables	Internet Search Data	Other Variables & Control Variables	Statistical Methods
Weng et al. (2018)	20 stocks (US market), 2013–2016	Stock prices	SVI for company names and related searches	Number of web page visitors (Wiki traffic), financial news sentiment, market indices and volume	Machine-learning ensemble methods
Xu et al. (2019)	Dow Jones Industrial, January 2004–July 2018	Volatility	Term 'stock' (US)	Macroeconomic fundamentals: gross domestic production, producer price index and industrial production	GARCH-MIDAS model

Notes: SVI = search volume index; BI = Baidu search volume index; VAR = vector autoregression; GARCH = generalised autoregressive conditional heteroskedasticity

3. METHODOLOGY

3.1 *Google Search Volume: A Proxy for Investor Attention*

Google Trends provides access to the relative online search volume for any query term submitted by Internet users to Google. Google search data has been available on a weekly and monthly basis since 2004. The actual search volume is normalised by the total number of searches for a specified region. Then, each search term is normalised by the maximum of searches. This scaling procedure makes Google search data conveniently presented in a [0, 100] interval, identified as the Google Search Volume Index (GSVI).

The GSVI increases when the actual number of searches increases compared with the average number of searches. Therefore, an increase does not necessarily imply a rise in the absolute number of online search queries. Increases mean primarily that the popularity of those particular query terms is increasing over time. In addition, due to the scaling procedure, the GSVIs of any two keywords are not comparable.

The search term used to identify a stock on Google is of crucial importance when using Google Trends. An investor who is searching for information regarding a specific company input either the firms' name or stock ticker symbol. As proposed by Bank et al. (2011) and Vlastakis and Markellos (2012), the present study used the GSVI of firm names rather than stock ticker symbols to capture the attention paid to particular stocks. First, we believe that the Portuguese retail investors who are the focus of this study are more likely to input a firm name to look for stock-specific information on Google. Second, search frequency results based on ticker symbols are lower in number compared with firm names, which would result in many missing values. Finally, we agree with Da et al.'s (2011) assertion that the GSVI for firm names incorporates some irrelevant components, such as individuals searching for company products or online support. However, for the present study, we assumed that these components are random noise.

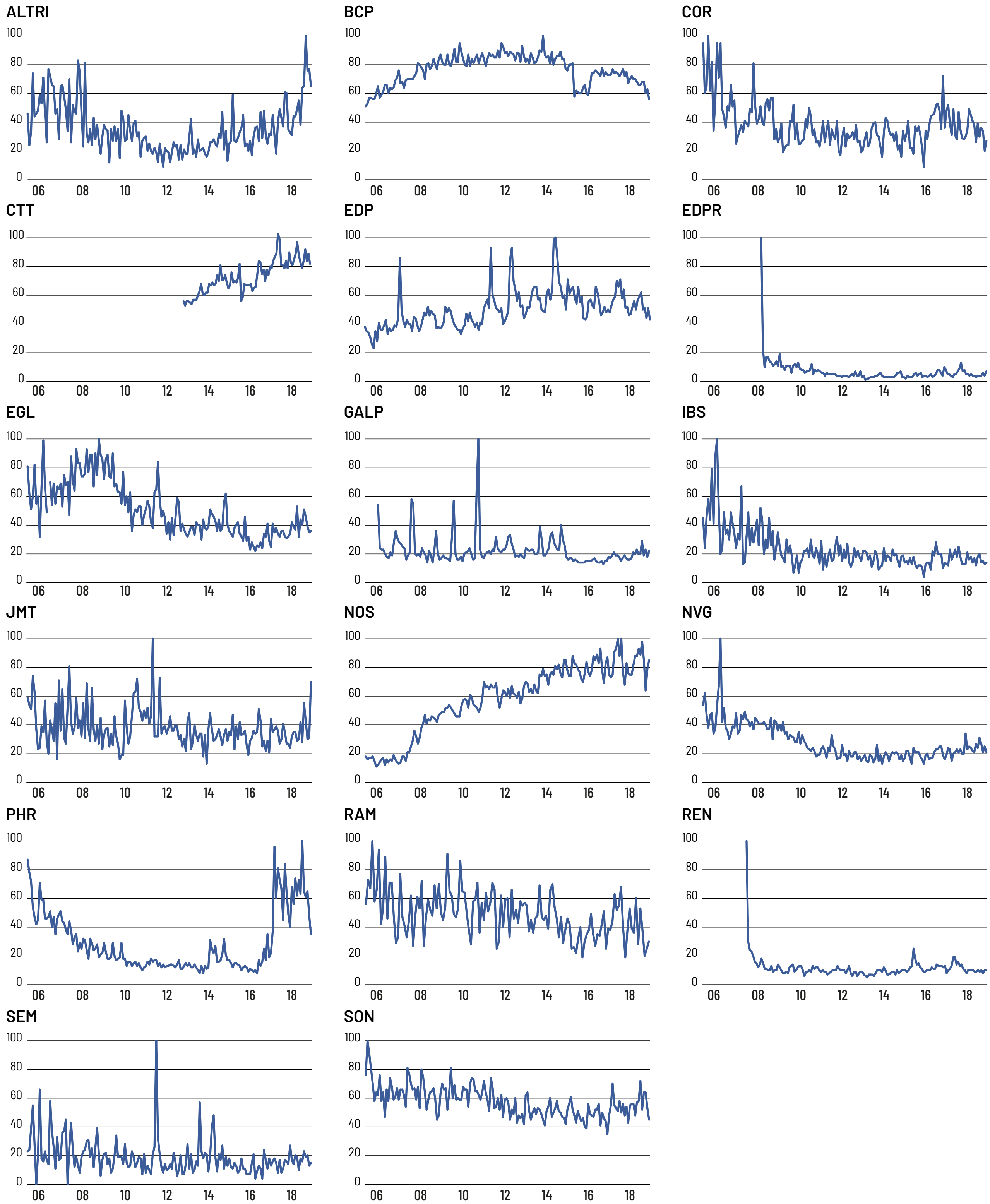
A regional filter based on the origin of queries determined from users' Internet Protocol addresses was added. Thus, all the search data in the current study originated in Portugal.

For each stock in our sample and the Portuguese Stock Index (PSI) 20, we manually drew the corresponding GSVI data. To address the possibility that company names could be searched in a variety of ways, we took two steps in order to identify the best queries for each company. First, we inserted the full company name and examined the related keyword terms offered by Google. We started searching by company name and retained all top searches, with a maximum of 10 related terms, to understand better how these words are employed by search engine users. Second, we picked the keyword with the highest search volume. We eliminated non-economic search terms (e.g. PT – Portugal) and search terms with too few valid GSVIs (e.g. several zero values). For those companies that registered corporate events involving changes in corporate brand names, a composite search was performed (e.g. 'Portugal Telecom + Pharol').

As we found many zero observations in the first two years (i.e. 2004 and 2005) for the PSI 20, we decided to start our analysis in March 2006. We began with the stocks listed in the PSI 20 as of February 2006 and traded in Euronext Lisbon. Based on the weekly search volume data, we opted to keep data from 2006 and 2014 on six companies – half of them from the financial sector. These were Banco Comercial Português (BCP), Banco Espírito Santo (BES), Banco Português de Investimento (BPI), Energias de Portugal (EDP), Portugal Telecom (PT) and Sonae Sociedade Gestora de Participações Sociais. This period allowed us to compare the behaviour of the variables under analysis before and after the economic crisis in 2008, for each company and the pool of six companies.

Next, we considered a sample of all the stocks listed in the PSI 20 as of September 2019 and traded in Euronext Lisbon (see Figure 3), for which Google search data were only available monthly. During the period under analysis (i.e. 2006–2019) various events were observed, such as delisting, new companies' entry and changes in corporate brand names.

Figure 3: Google search volume for stocks listed in PSI



Based on the press releases published on the company's websites, we were able to identify a variety of corporate events that might have triggered top volume searches, namely, the maximum possible SVI (i.e. 100). These corporate events include notifications to stakeholders that the dividend for the previous financial year will be paid (e.g. Altri,[1] May 2019), communication of capital increases (e.g. BCP,[2] January 2017) and notifications of new operational plans (e.g. CTT,[3] December 2017). Other relevant events are announcements of financial results (e.g. Galp,[4] July 2011; Pharol,[5] March 2019), communication of listing in stock markets (e.g. EDP Renovaveis [EDPR]), communication of rating by Moody's (e.g. EDP,[6] February 2015) and sales and market share reports (e.g. Jeronimo Martins,[7] January 2012).

To address our concerns about outliers, each series was winsorised at the 5% level (i.e. 2.5% in each tail). We followed Da et al.'s (2011) approach because this offers a simpler interpretation. Thus, we took the logarithm of GSVIs, denoted by SVI for each stock *i* for week *t*:

$$SVI_{i,t} = ln(GSVI_{i,t})$$

We also expressed market level Internet searches (i.e. for PSI 20) as:

$$SVI_{M,t} = ln(GSVI_{M,t})$$

1 See http://www.altri.pt/~/media/Files/A/Altri-V2/press-release/2019/ALTRI20190529EN.pdf.

2 See https://ind.millenniumbcp.pt/en/Particulares/_layouts/BCP.SDC.FEP.Foundation.Presentation/Error/404.aspx.

3 See https://www.ctt.pt/contentAsset/raw-data/01ff5161-c64e-47c0-b426-650012e61d97/ficheiro/edd10fa0-8887-4d-26-ac2b-2ba1d2e3cc35/export/CTT%20Operational%20Transformation%20Plan_PT_FINAL.pdf.

4 See https://www.galp.com/corp/Portals/0/Recursos/Investidores/SharedResources/Resultados/EN/2011-2T-RT/ACF/1H2011RESULTS.pdf.

5 See http://pharol.pt/en-us/press-releases/comunicados-legais/Pages/2019/Com06Mar19_PH_AG.aspx.

6 See https://www.edp.com/sites/default/files/portal.com/documents/Moodys%20Upgrade_EN_20150213.pdf.

7 See https://www.jeronimomartins.com/wp-content/uploads/com/2012/en/com_20120110_1_en.pdf.

3.2 *Stock Market Activity*

This section presents the measures of stock market activity used in this study: volume, returns and volatility. If $P_{i,t}$ is the observed weekly closing price of stock $_i$, then weekly changes in price and returns are denoted by $r_{i,t}$:

$$r_{i,t} = ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right)$$

This study also investigated the association between investor attention and market activity from the perspective of historical volatility. Realised volatility is one of the most popular measures of historical volatility. We followed the approach used in previous research on stock market volatility (Dimpfl and Jank, 2011; Vlastakis and Markellos, 2012) and proxy volatility by the standard deviation of returns. The realised volatility for stock $_i$ for week t (RVit) is computed from daily data, where $r_{i,j}^2$ corresponds to the squared return of the *i*th stock for day j:

$$RV_{i,t} = \sqrt{\sum_{j=1}^{5} r_{i,j}^2}$$

The proxy for firm size was the logarithm of stock volume i for week t, denoted by $\boldsymbol{vol}_{i,t}$.


Google

4. RESULTS

4.1 Descriptive Statistics for Weekly Data

Table 2 presents the descriptive statistics for GSVIs. Some variability exists in Portuguese investor attention across the six stocks studied. BCP has the highest GSVI (mean = 58.66) and PT the lowest (mean = 30.82) in this sample. The normality distribution of GSVIs was rejected in all cases. The variables referred to hereafter were logarithmically transformed.

Table 2: Descriptive statistics of Google search volume

Variable	Mean	Median	Standard Deviation	Minimum	Maximum	Skewness	Kurtosis	Kolmogorov--Smirnov Test	
GSVI_BCP	58.66	55	12.70	38	100	1.13	0.61	0.15	***
GSVI_BES	58.36	61	11.89	27	100	-0.55	0.06	0.10	***
GSVI_BPI	51.00	47	14.53	28	100	1.40	1.23	0.20	***
GSVI_EDP	34.31	32	11.17	16	100	2.48	10.32	0.13	***
GSVI_PT	30.82	23	19.43	7	100	1.28	0.88	0.17	***
GSVI_SONAE	57.63	57	12.17	29	100	0.52	0.60	0.06	***

Note: *, ** and *** denote 10%, 5% and 1% significance levels, respectively.

SVI is stationary around a deterministic trend for all the stocks selected. The next table displays the results of two unit root tests of Google search names: the augmented Dickey-Fuller (ADF) and Philips-Perron (PP). Therefore, a trend variable was included in the regression, as was done in Aouadi et al.'s (2013) study.

Table 3: Unit root test on Google search volume

Stock	ADF	PP
BCP	-4.43***	-6.29***
BES	-9.17***	-4.76***
BPI	-3.83**	-11.47***
EDP	-8.62***	-8.93***
PTC	-6.93***	-15.81***
SON	-7.73***	-15.43***

Notes: The null hypothesis is the existence of a unit root (i.e. stationarity); *, ** and *** denote 10%, 5% and 1% significance levels, respectively.

4.2 *Google Search Volume and Stock Market Volatility for Weekly Data*

The following regression was formulated in order to analyse the relationship between information attention and realised volatility:

$$RV_{i,t=}\alpha + \gamma_1 SVI_{i,t}+\gamma_2 SVI_{M,t} + \beta_1 vol_{i,t} + \beta_2 r_{i,t} + \beta_3 RV_{i,t-1} + \beta_4 t + \varepsilon_{i,t}$$

in which $SVI_{i,t}$ is the information attention for stock i for week t and $SVI_{M,t}$ is market-related information attention (i.e. Google search volume for the term PSI 20). The control variables are as follows: $vol_{i,t}$ is the logarithm of market volume, $r_{i,t}$ is stock market return, $RV_{i,t-1}$ is one lag of stock market volatility and t is a time trend.

For three of the six stocks studied, the stock-specific SVI variable was significant at the 5% level, but with mixed signals. For BCP and BES, investor attention appears to increase in volatility by incorporating more information into prices, but for BPI, attention reduces volatility, possibly by reducing uncertainty.

Market-related SVI is significantly positive for five of the six stocks. Moreover, for financial stocks only, the effect of the stock-specific SVI variable is stronger than that of market-related SVI. Of all the non-financial stocks, only the SVI market-related variable is significant. According to Peng and Xiong (2006), this could be explained as investors' tendency to process more market than stock specific information.

Table 5: Model estimates

Parameter stimates	BCP	BES	BPI	EDP	PTC	SON	Pooled Sample
α	-0.0227***	-0.0227***	0.0062**	-0.0038**	-0.0156***	-0.0227***	-0.0074***
	(0.0056)	(0.0049)	(0.0030)	(0.0018)	(0.0047)	(0.0056)	(0.0009)
y_1	0.0045***	0.0061***	-0.0036***	0.0000	0.0019*	0.0045***	0.0006***
	(0.0012)	(0.0013)	(0.0009)	(0.0005)	(0.0010)	(0.0012)	(0.0002)
y_2	0.0018***	0.0004	0.0016***	0.0009***	0.0019***	0.0018***	0.0009***
	(0.0006)	(0.0004)	(0.0005)	(0.0003)	(0.0006)	(0.0006)	(0.0002)
β_1	0.0013***	0.0067***	0.0039***	0.00169***	0.0061***	0.0013***	0.0068***
	(0.0001)	(0.0008)	(0.0004)	(0.0003)	(0.0008)	(0.0001)	(0.001)
β_2	-0.0017	0.0024	0.0096***	-0.0061**	0.0079**	-0.0017	0.0006***
	(0.0039)	(0.0030)	(0.0034)	(0.0031)	(0.0040)	(0.0039)	(0.0001)
β_3	0.0315	0.2036***	0.0818**	0.5548***	0.2058***	0.0315	0.2283***
	(0.0433)	(0,0438)	(0.0374)	(0.0405)	(0.0448)	(0.0433)	(0.0188)
β_4	-0.0004***	-0.0007***	0.0004***	0.0001	0.0003**	-0.0004***	0.0001
	(0.0001)	(0.0002)	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Adj.R^2	0.3148	0.3414	0.4740	0.3837	0.2662	0.3148	0.1978

Notes: *, ** and *** denote 10%, 5% and 1% significance levels, respectively; Newey-West heteroskedasticity and autocorrelation consistent (HAC) standard errors and covariance are employed in the estimation of cross-section models; $RV_{i,t} = \alpha + y1\, SVI_{i,t} + y_2\, SVI_{M,t} + \beta_1\, vol_{i,t} + \beta_2\, r_{i,t} + \beta_3\, RV_{i,t-1} + \beta_4\, t + \sum_{i,t}$, where $SVI_{i,t}$ is the information attention for stock i for week *t*, $SVI_{M,t}$ is the Google search volume for the term PSI 20, $vol_{i,t}$ is the logarithm of market volume, $r_{i,t}$ is stock market return, $RV_{i,t-1}$ is one lag of stock market volatility and t is a time trend.

4.3 *Effect of Market States*

We next shifted our attention to the possible impact of market states on the influence between information demand and stock market activity. In line with Vlastakis and Markellos's (2012) work, a dummy variable was defined. This variable took the value of one when a large market price change occurs (i.e. weeks for which the deviation of the absolute return of the market from its mean is more than one standard deviation) and zero in all other cases that follow the opposite pattern. All other weeks were considered to be a 'low return state'. More specifically, the high-return state dummy variable was defined as:

$$H_t = \begin{cases} 1\ if\ \left(abs\_r_{M,t} - \overline{abs\_r_M}\right) > \sigma\_abs\_r_M \\ 0\ otherwise \end{cases}$$

in which $\boldsymbol{abs\_r_{M,t}}$ is the absolute market return for week t, $\boldsymbol{abs\_r_M}$ is the average absolute market return over the complete sample and $\boldsymbol{\sigma\_abs\_r_M}$ is the standard deviation of absolute market return over the complete sample. We then considered an interaction term between stock-specific investor attention and the market state variable.

Model 2 was thus defined as follows:

$$RV_{i,t}=\alpha + \gamma_1 SVI_{i,t} + \gamma_2 SVI_{i,t} \times H_t + \gamma_3 SVI_{M,t} + \beta_1 vol_{i,t} + \beta_2 r_{i,t} + \beta_3 RV_{i,t-1} + \beta_4 t + \varepsilon_{i,t}$$

in which all other variables are defined as previously. In addition, $\boldsymbol{SVI_{i,t}}$ is the information attention for stock i for week t, $\boldsymbol{SVI_{M,t}}$ is market-related information attention, $\boldsymbol{vol_{i,t}}$ is the logarithm of market volume, $\boldsymbol{r_{i,t}}$ is stock market return, $\boldsymbol{RV_{i,t-1}}$ is one lag of stock market volatility and t is a time trend.

Those stocks that register a significantly positive stock-specific SVI in Model 1 present a significantly positive stock-specific SVI for the high-return market state. SON (i.e. SONAE) stocks, whose SVI coefficient in Model 1 are only statistically significant at the 10% level, also have the same profile. Thus, the results support the conclusion that SVI's impact on realised volatility is stronger for high-return market states. These findings are in accordance with those reported by Vlastakis and Markellos (2012).

Table 5: Market states model estimates

Parameter stimates	BCP	BES	BPI	EDP	PTC	SON	Pooled Sample
α	-0.0200***	-0.0178	0.0070**	-0.0032*	-0.0021***	-0.0071***	-0.0058
	(0.0056)	(0.0047)	(0.0030)	(0.0017)	(0.0026)	(0.0017)	(0.0009)
γ_1	0.0040***	0.0050***	-0.0035***	0.0002	-0.0005	0.0014**	0.0005***
	(0.0012)	(0.0013)	(0.0009)	(0.0001)	(0.0006)	(0.0006)	(0.0002)
γ_2	0.0006***	0.0008***	0.0002	0.0003***	0.0006***	0.0008***	0.0008***
	(0.0001)	(0.0001)	(0.0002)	(0.0000)	(0.0001)	(0.0001)	(0.0000)
γ_3	0.0014**	0.0007	0.0006**	0.0008***	0.0012***	0.0018***	0.0014***
	(0.0006)	(0.0005)	(0.0001)	(0.0003)	(0.0004)	(0.0006)	(0.0002)
β_1	0.0012***	0.0054***	0.0038***	0,0016***	0,0058***	0.0059***	0.0067***
	(0.0001)	(0.0007)	(0.0002)	(0,0004)	(0,0001)	(0.0008)	(0.0005)
β_2	-0.0009	0.0039	0.0012***	-0.0047*	-0.0041	0.0102**	0.0085***
	(0.0040)	(0.0029)	(0.0005)	(0.0031)	(0.0033)	(0.0039)	(0.0016)
β_3	0.0336	0.1979***	0.0884***	0.5459***	0.2413***	0.1791***	0.2117***
	(0.0432)	(0.0417)	(0.0366)	(0.0403)	(0.0464)	(0.0435)	(0.0186)
β_4	-0.0003***	-0.0005***	0.0004***	0.0000	-0.0000	0.0002***	0.0008
	(0.0001)	(0.0006)	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0007)
Adj.R^2	0.3159	0.4035	0.4994	0.4053	0.2348	0.3170	0.2530

Notes: *, ** and *** denote 10%, 5% and 1% significance levels, respectively; Newey-West HAC standard errors and covariance are employed in the estimation; $RV_{i,t} = \alpha + \gamma_1 SVI_{i,t} + \gamma_2 SVI_{M,t} + \beta_1 vol_{i,t} + \beta_2 r_{i,t} + \beta_3 RV_{i,t-1} + \beta_4 t + \sum_{i,t}$, in which $SVI_{i,t}$ is the information attention for stock i for week t, $SVI_{M,t}$ is market-related information attention, $vol_{i,t}$ is the logarithm of market volume, $r_{i,t}$ is stock market return, $RV_{i,t-1}$ is one lag of stock market volatility, t is a time trend and H_t is the high-return state dummy variable.

4.4 *Financial Crisis Effect*

The sample period, March 2006 to February 2014, includes a period of economic downturn. We thus had to check further Model 1's stability, as defined previously. To address this issue, a Quandt-Andrews (QA) breakpoint test was performed. QA test results provide the maximum likelihood ratio F-statistics for each regression under the null hypothesis of 'no structural break points'. This stability test detected one or more structural breakpoints in the sample, with a chosen trimming region of 15% of the sample period. We included $SVI_{i,t}$ as varying regressors. The hypothesis of a stable link between market volatility and Google search volume for companies' names was rejected for all the stocks analysed (see Table 7).

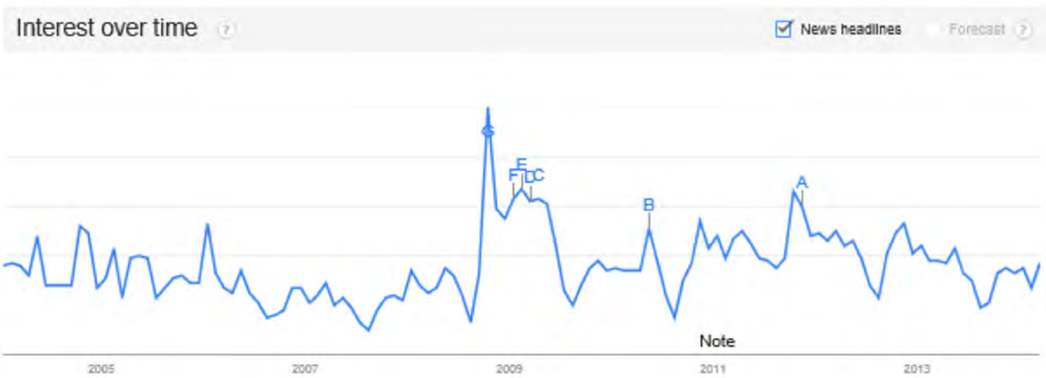
Table 6: Results of QA breakpoint tests

	BCP	BES	BPI	EDP	PTC	SON
QA	58.51***	78.52***	23.30***	26.47***	27.022***	16.27***

Note: *, ** and *** denote 10%, 5% and 1% significance levels, respectively.

Subsequently, we adopted the approach proposed by Vlastakis et al. (2012) in order to examine the stability of Model 2's results. We tested the existence of differences in investor attention's impact on stock volatility by splitting the available sample into two parts. The second sub-sample contains the recent financial crisis. To decide how to split our sample, we analysed Google search data for Portugal with the word 'crisis' as a key term. Based on the results presented in Figure 4, two periods were defined as after and before October 2008.

Figure 4: Google search volume for search query 'crisis'



Source: Google Trends (n.d.)

A second stability test – Chow's breakpoint test – confirmed that a shift in the market state has a significant effect on the relationship between stock-specific SVI and realised volatility of all stocks.

Table 7: Results of Chow's breakpoint test

	BCP	BES	BPI	EDP	PTC	SON
F	14.33***	12.44***	4.01**	16.86***	15.89***	4.96**

Note: *, ** and *** denote 10%, 5% and 1% significance levels, respectively.

Therefore, we decided to estimate a third model by adding a dummy variable to account for the economic crisis effect, splitting our sample into the two aforementioned periods. To explore further the economic crisis's impact on stock specific SVI model estimates, a regression framework using a dummy variable was employed. The dummy variable ***dCt*** allowed us to evaluate any differences between the two periods in the impact of stock-specific information demand on market activity.

Model 3 was defined as follows:

$$RV_{i,t=}\alpha + \gamma_1 SVI_{i,t} + \gamma_2 SVI_{i,t} \times H_t + \gamma_3 SVI_{i,t} \times dC_t + \gamma_4 SVI_{M,t} \times dC_t + \beta_1 vol_{i,t} + \beta_2 ret_{i,t} + \beta_3 RV_{i,t-1+}\beta_4 t + \varepsilon_{i,t}$$

in which $SVI_{i,t}$ is the information attention for stock i for week t and $SVI_{M,t}$ is market-related information attention. In addition, $vol_{i,t}$ is the logarithm of market volume, $r_{i,t}$ is stock market return, $RV_{i,t-1}$ is one lag of stock market volatility, t is a time trend, H_t is the market state dummy variable and dC_t is the dummy crisis variable. As the interaction term between stocks' specific SVI and the economic crisis dummy variable is positive and statistically significant for five of the six cases, the impact of investor attention appears to be stronger during the economic crisis period.

Table 8: Economic crisis model estimates

Parameter stimates	BCP	BES	BPI	EDP	PTC	SON	Pooled Sample
α	-0.0179***	-0.0177***	0.0064**	0.0042**	-0.0050**	-0.0131***	-0.0053***
	(0.0058)	(0.0050)	(0.0031)	(0.0017)	(0.0026)	(0.0046)	0.0009
y_1	0.0033**	0.0047***	-0.0036***	0.0000	-0.0000	0.0012	0.0004**
	(0.0013)	(0.0014)	(0.0009)	(0.0005)	(0.0005)	(0.0010)	0.0002
y_2	0.0007***	0.0007***	0.0005***	0.0006***	0.0007***	0.0006***	0.0008***
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0002)	0.0000
y_3	0.0019***	0.0007	0.0016***	0.0011***	0.0016***	0.0020***	0.0003***
	(0.0006)	(0.0005)	(0.0006)	(0.0003)	(0.0004)	(0.0005)	(0.0000)
y_4	0.0019***	0.0007	0.0016***	0.0011***	0.0016***	0.0020***	0.0015***
	(0.0006)	(0.0005)	(0.0006)	(0.0003)	(0.0004)	(0.0005)	(0.0002)
β_1	0.0014***	0.0076***	0.0040***	0.0023***	0.0069***	0.0068***	0.0068***
	(0,0001)	(0.0008)	(0.0002)	(0.0004)	(0.0001)	(0.0009)	(0.0009)
β_2	-0.0020	0.0021	0.0096***	-0.0060**	-0.0055*	0.0071*	0.0085***
	(0.0039)	(0.0029)	(0.0034)	(0.0030)	(0.0033)	(0.0039)	(0.0015)
β_3	0.0236	0.1867***	0.07926**	0.5087***	0.1919***	0.1746***	0.2054***
	(0.0429)	(0.0433)	(0.03721)	(0.0411)	(0.0463)	(0.0448)	(0.0002)
β_4	-0.0009***	-0.0012***	0.0001***	-0.0003**	-0.0003**	-0.0001	-0.0002***
	(0.0002)	(0.0002)	(0.000)	(0.0001)	(0.0001)	(0.0001)	(0.0000)
Adj.R^2	0.3291	0.3627	0.4825	0.40834	0.2351	0.2901	0.2664

Notes: *, ** and *** denote 10%, 5% and 1% significance levels, respectively; Newey-West HAC standard errors and covariance are employed in the estimation; $RV_{i,t} = \alpha + y1\, SVI_{i,t} + y_2\, SVI_{it} + H_t + y_3\, SVI_{i,t} + dC_t\, \beta_2 + y_3\, SVI_{M,t} \times dC_t + \beta_1 vol_{it} + \beta_2 ret_{i,t} + \beta_3 ret_{i,t-1} + \beta_4 ret_{it} + \sum_{i,t}$, in which $SVI_{i,t}$ is the information attention for stock i for week t, $SVI_{M,t}$ is market-related information attention, $vol_{i,t}$ is the logarithm of market volume, $r_{i,t}$ is stock market return, $RV_{i,t-1}$ is one lag of stock market volatility, t is a time trend and H_t is the high-return state dummy variable and C_t is the crisis dummy variable.

4.5 *Google Search Volume and Stock Market Volatility for Monthly Data*

To test the results' robustness, the model presented in section 4.2 was tested for monthly data on the 17 companies presented in Figure 3 above. For the seven stocks under analysis, the stock-specific SVI variable is significant at the 5% level. In addition, as these estimates are positive, the conclusion can be drawn that investor attention appears to increase volatility by incorporating more information into prices for the seven stocks. Market-related SVI is significantly positive for 13 of the 17 stocks analysed.

The results obtained for monthly data are in accordance with those for the weekly data, as the data on statistically significant market-related SVI show this is higher in volume than the stock specific SVI is. The model estimates indicate that Portuguese investors may tend to consider more information about markets than specific stocks. The estimates for both market-based and stock-related Google search variables suggest a positive relationship between Internet search volume and stock market volatility.

Table 9: Model estimates

	α	γ_1	γ_2	β_1	β_2	β_3	β_4	Adj.R^2
ALTR	-0.5149**	0.0011**	0.0023***	0.0397**	-0.1128	0.0033***	-0.0006*	0.332
	(0.2562)	(0.0005)	(0.0008)	(0.0164)	(0.0723)	(0.0007)	(0.0004)	
BCP	-2.0297***	0.0039***	0.0024**	0.1048***	-0.3417***	0.0048***	-0.0024***	0.532
	(0.416)	(0.0011)	(0.001)	(0.023)	(0.0846)	(0.0006)	(0.0005)	
COR	0.1811* (0.1051)	-0.0004	0.0017**	-0.0035	0.1136	0.0057***	-0.0009***	0.371
	-0.9789**	(0.0006)	(0.0007)	(0.0078)	(0.0982)	(0.0007)	(0.0003)	
CTT	(0.3739)	0.0060***	0.0016 (0.001)	0.0776***	-0.5676***	0.0021**	-0.0018**	0.395
	-0.7500**	(0.0015)	0.0018***	(0.0235)	(0.1219)	(0.001)	(0.0008)	
EDP	(0.338)	0.0004	(0.0006)	0.0408**	-0.2373**	0.0066***	-0.0005**	0.512
	-1.0652	(0.0005)	0.0019***	(0.0174)	(0.0937)	(0.0006)	(0.0003)	
EDPR	(0.2565)	0.0032	(0.0003)	0.0644***	-0.2369***	0.0065***	-0.0001	0.716
	-1.162***	(0.0021)	-0.1198**	(0.0143)	(0.0796)	(0.0006)	(0.0004)	
EGL	(0.2641)	0.0929	(0.0505)	0.0036***	-0.0001	0.0014***	0.0003	0.448
	-3.1457***	(0.0177)	0.0014**	(0.0006)	(0.0004)	(0.0009)	(0.0007)	
GALP	(0.4208)	0.0008	(0.0006)	0.1893***	-0.181**	0.0054***	-0.0006**	0.715
	0.0218***	(0.0006)	0.0008**	(0.0248)	(0.0711)	(0.0005)	(0.0002)	
IBS	(0.0748)	0.0012**	(0.0005)	0.0080	-0.0771 (0.053)	0.007***	-0.0005**	0.551
	-2.2028**	(0.0005)	0.0015**	(0.0059)	-0.095	(0.0006)	(0.0002)	
JMT	(0.3258)	0.0097***	(0.0006)	0.1376***	(0.0784)	0.0042***	-0.0006**	0.631
	-0.8624***	(0.0006)	0.0005	(0.0199)	-0.1014	(0.0006)	(0.0002)	
NOS	(0.1774)	0.0005	(0.0006)	0.0616***	(0.0727)	0.0054***	-0.0008*	0.542
	-0.2562**	(0.0007)	0.2163**	(0.0118)	-0.0004***	(0.0006)	(0.0004)	
NVG	(0.1096)	0.0232	(0.0712)	-0.0006***	(0.0002)	0.0011***	0.0049*	0.395
	-3.377***	(0.0072)	0.0007	(0.0007)	0.0275	(0.0005)	(0.0006)	
PHR	(0.4344)	-0.0009	(0.0012)	0.1927***	(0.0863)	0.0031	0.0012***	0.593
	0.3470**	(0.0006)	-0.0004	(0.0242)	-0.2466**	(0.0006)	(0.0005)	
RAM	(0.141)	0.0001	(0.0011)	-0.0094	(0.1199)	0.0063***	-0.0006	0.523
	-0.2288**	(0.0009)	0.0019***	(0.0109)	-0.3643***	(0.0007)	(0.0005)	
RENE	(0.1118)	0.0032**	(0.0005)	0.0235***	(0.1039)	0.0047***	-0.0012***	0.510
	-0.6494***	(0.0018)	0.0009**	(0.0082)	-0.0476	(0.0006)	(0.0002)	

Table 9: Model estimates

	α	y_1	y_2	β_1	β_2	β_3	β_4	Adj.R^2
SEM	-0.6494***	0.0004	0.0009**	0.047***	-0.0476	0.0059***	0.0004*	0.471
	(0.1493)	(0.0004)	(0.0004)	(0.0098)	(0.0628)	(0.0006)	(0.0002)	
SON	-0.3272	-0.0264**	0.317***	0.0048**	-0.0007***	0.0020***	-0.0009**	0.427
	(0.2589)	(0.0041)	(0.0779)	(0.0006)	(0.0003)	(0.0007)	(0.001)	

Notes: *, ** and *** denote 10%, 5% and 1% significance levels, respectively; Newey-West HAC standard errors and covariance were used to estimate cross--section models; $RV_{i,t} = \alpha + y1\, SVI_{i,t} + y_2\, SVI_{MT} + \beta_1 vol_{it} + \beta_2 r_{i,t} + \beta_3 RV_{i,t-1} + \beta_4 t + \Sigma_{i,t}$, in which $SVI_{i,t}$ is the information attention for stock i for week t, $SVI_{M,t}$ is the Google search volume for the term 'PSI 20', $vol_{i,t}$ is the logarithm of market volume, $r_{i,t}$ is stock market return, $RV_{i,t-1}$ is one lag of stock market volatility and t is a time trend.

5. CONCLUSION

This study examined investor stock-specific and market-related attention and its relationship to stock market volatility. As a proxy for investor attention, we used a measure based on Internet search volume for the keywords of stocks traded on Euronext Lisbon. As reported in previous empirical studies, Google search volume is a reliable proxy for investor attention. Moreover, the model estimates for the selected stocks indicate that Google search volume is a significant determinant of contemporaneous stock market historical volatility. The effects are robust even after controlling for variations in market return and market volume.

In a second step, we tested whether the influence of stock-specific Google search data on realised volatility varies according to market states. According to the model estimates, investor attention's impact appears to be more sensitive to a high-return market state. This result is in accordance with the findings provided by Vlastakis and Markellos (2012), who reached the same conclusion for the largest stocks traded on the NYSE and NASDAQ.

As the first sample period under analyses (i.e. March 2006 to February 2014) includes a period of economic downturn, we performed additional stability tests by splitting the data into two periods. The results indicate that the impact of Google search data on realised volatility becomes stronger during periods of crisis. This finding, however, is not in line with Aouadi et al.'s (2013) results for the French stock market. The cited authors report stability in the model estimates after controlling for any economic crises. Thus, future studies of this question are needed to ensure the present findings' external validity.

The last regression analysis covered monthly data and the companies listed in the PSI 20 in September 2019. The results remain robust, revealing the importance of Google Search data as a tool to explain market volatility.

These findings contribute to a better understanding of Portuguese stock market activity. As the GSVI is a reliable proxy for investor attention, future studies could analyse investor attention's impact on different variables of stock market activity, such as returns, abnormal returns, volume, abnormal volume and illiquidity. Future studies could also look into the forecasting capabilities of GSVI data. Finally, checking whether investor attention can be used as an indicator of systemic risk in the market could provide meaningful results for regulators and market participants.

REFERENCES

- Aouadi, A., Aroudi, M. and Teulon, F. (2013). Investor attention and stock market activity: evidence from France. *Economic Modelling*, 35, 674–681.

- Bank, M., Larch, M. and Peter, G. (2011). Google search volume and its influence on liquidity and returns of German stocks. *Financial Markets and Portfolio Management*, 25 (3), 239–264.

- Barber, B. M. and Odean, T. (2008). All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, 21, 785–818.

- Chemmanur, T. and Yan, A. (2009). Advertising, attention, and stock returns (Working paper). Boston College, Chesnut Hill, MA, and Fordham University, New York, NY, USA.

- Chouliaras, A. and Grammatikos, T. (2013). News flow, web attention and extreme returns in the European financial crisis. *MPRA Paper 51335*. University Library of Munich, Munich, Germany.

- Da, Z, Engelberg, J. and Gao, P. (2014). The sum of all FEARS: investor sentiment and asset prices. *Review of Financial Studies*, 28 (1), 1–32.

- Da, Z., Engelberg, J. and Gao, P. (2010). In search of fundamentals (Working paper). University of Notre Dame, Notre Dame, IL, and University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

- Da, Z., Engelberg, J. and Gao, P. (2011). In search of attention. *The Journal of Finance*, 66 (5), 1461–1499.

- Dellavigna, S. and Pollet, J. M. (2009). Investor attention and Friday earnings announcements. *Journal of Finance*, 64, 709–749.

- Dergiades, T., Milas, C. and Penagiotidis, T. (2013). Tweets, Google Trends and sovereign spreads in the GIIPS. *GreeSE Paper No. 78*. London School of Economics and Political Science, London, UK.

REFERENCES

- Dimpfl, T. and Jank, S. (2011). Can internet search queries help to predict stock market volatility? *CFR Working Papers, 11-15*. University of Cologne, Centre for Financial Research (CFR), Cologne, Germany.
- Drake, M. S., Darren, R. T. and Thornock, J. R. (2012). Investor information demand: evidence from Google searches around earnings announcements, *Journal of Accounting Research*, 50 (4), 1001-1040.
- Dzielinski, M. (2012). Measuring economic uncertainty and its impact on the stock market, *Finance Research Letters*, 9, 167-175.
- Fang, L. and Peress, J. (2009). Media coverage and the cross-section of stock returns. *Journal of Finance*, 64, 2023-2052.
- Frieder, L. and Subrahmanyam A. (2005). Brand perceptions and the market for common stock. *The Journal of Financial and Quantitative Analysis*, 40 (1), 57-85.
- Google Trends. (n.d.). Available at: https://www.google.com/trends/ (accessed 15 December 2019).
- Grullon, G., Kanatas, G. and Weston, P. J. (2004). Advertising, breath ownership, and liquidity. *Review of Financial Studies*, 17, 439-461.
- Heiberger, R. H. (2015). Collective attention and stock prices: evidence from google trends data on Standard and Poor's 100. PLoS ONE, 10, e0135311.
- Hirshleifer, D. and Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics*, 36 (1-3), 337-386.
- Latoeiro, P. and Ramos, S. (2013). O Google como barómetro da atenção do investidor [Google as an investor attention barometer]. *Caderno do Mercado de Valores Mobiliários*, 46 (4), 3-36.

REFERENCES

- Lobão, J., Pacheco, L. and Pereira, C. (2017). The use of the recognition heuristic as an investment strategy in European stock markets. *Journal of Economics*, Finance and Administrative Science, 22, 207–223.
- Merton, R. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, 42, 483–510.
- Mondria, J. and Quintana-Domeque, C. (2012). Financial contagion and attention allocation. *The Economic Journal*, 123 (568), 429–454.
- Mondria, J., Wu, T. and Zhang, Y. (2010). The determinants of international investment and attention allocation. *Journal of International Economics*, 82, 85–95.
- Peng, L. and Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, 80, 563–602.
- Seasholes, M. S. and Wu, G. (2007). Predictable behavior, profits, and attention. *Journal of Empirical Finance*, 14, 590–610.
- Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50 (3), 665–690.
- Smith, G. P. (2012). Google Internet search activity and volatility prediction in the market for foreign currency. *Finance Research Letters*, 9 (2), 103–110.
- Takeda, F. and Wakao, T. (2014). Google search intensity and its relationship with returns and trading volume of Japanese stocks. *Pacific-Basin Finance Journal*, 27, 1–18.
- Vlastakis, N. and Markellos, R. N. (2012). Information search and stock market volatility. *Journal of Banking and Finance*, 36 (6), 1808–1821.
- Vozlyublennaia, N. (2014). Investor attention, index performance, and return predictability. *Journal of Banking & Finance*, 41, 17–35.

REFERENCES

- Weng, B., Lu, L., Wang, X., Megahed, F. M. and Martinez, W. (2018). Predicting short-term stock prices using ensemble methods and online data sources. *Expert Systems with Applications*, 112, 258–273.

- Xu, Q., Bo, Z., Jiang, C. and Liu, Y. (2019). Does Google search index really help predicting stock market volatility? Evidence from a modified mixed data sampling model on volatility. *Knowledge-Based Systems*, 166, 170–185.

- Yuan, Y. (2011). Attention and trading (Working paper). Wharton School of the University of Pennsylvania, Philadelphia, PA, USA.

- Zhang, W., Shen, D., Zhang, Y. and Xiong, X. (2013). Open source information, investor attention and asset pricing. *Economic Modelling*, 33, 613–619.