Journal Pre-proof

Google Search-Based Sentiment Indexes

Ana Oliveira-Brochado

PII: S0970-3896(17)30271-9
DOI: https://doi.org/10.1016/j.iimb.2019.10.015
Reference: IIMB 368

To appear in: *IIMB Management Review*

Received date: 1 June 2017
Revised date: 23 February 2018
Accepted date: 17 October 2019

Please cite this article as: Ana Oliveira-Brochado , Google Search-Based Sentiment Indexes, *IIMB Management Review* (2019), doi: https://doi.org/10.1016/j.iimb.2019.10.015

GOOGLE SEARCH-BASED SENTIMENT INDEXES

Ana Oliveira-Brochado

Instituto Universitário de Lisboa, ISCTE Business School, DINÂMIA'CET – IUL, Centre for Socioeconomic and Territorial Studies, Business Research Unit (BRU-IUL)

Av. Forças Armadas, 1649-026 Lisboa, Portugal. Tel.+351.2104640 14

ana.brochado@iscte-iul.pt

ABSTRACT

This study sought to confirm whether Internet search-based data have the potential both to reveal populations' underlying beliefs directly and to affect stock markets of countries – in this case, Portugal. Based on the Internet search volume of several queries related to household concerns, we constructed two Google-based sentiment measures encompassing both positive and negative search terms. The results reveal that both measures are correlated with aggregate stock market returns, trading volume and abnormal trading volume. The results contribute to the literature by highlighting the different roles of positive and negative sentiment in stock market activity.

SUMMARY

As Internet searches are generated by agents' spontaneous behaviour, these search patterns may possess signalling properties. This study sought to confirm whether Internet search-based data have the potential both to reveal populations' underlying beliefs directly and to affect stock markets of countries – in this case, Portugal. Based on the Internet search volume of several queries related to household concerns, we constructed two Google-based sentiment measures encompassing both positive and negative search terms (i.e. a Positive Sentiment Index [PSI] and a Negative Sentiment Index [NSI]).

The final list of negative terms comprised ‘austerity’, ‘taxes’, ‘rents,’ ‘Euribor’, ‘crisis’, ‘debt’, ‘finance’, ‘gold price’, ‘unemployment’ and ‘poverty’. The final list of positive terms included ‘stocks’, ‘consume’, ‘credit’, ‘GDP’, ‘Lisbon stock market’, ‘dividends’, ‘profits’, ‘investment’, ‘entrepreneurship’ and ‘partnership’.

The results reveal that both measures are correlated with aggregate stock market returns, trading volume and abnormal trading volume. We also found that positive sentiment has a stronger impact on these stock market variables than negative sentiment does. In addition, the results show that the proposed sentiment measures are significantly useful when making short-term predictions of market returns and volume. These findings confirm the validity of using sentiment measures based on search queries since they are available for shorter periods than other well-known measures of market sentiment are. The results contribute to the literature by highlighting the different roles of positive and negative sentiment in stock market activity.

KEYWORDS: investor sentiment, search-based data, stock market, Portugal

1. INTRODUCTION

Keynes (1936) maintained that investors' 'animal spirits' could be used to justify wild movements in stock market prices seemingly unjustified by fundamental principles. Some 50 years later, other authors further elaborated on the role of investor sentiment in stock market activity (e.g. Barberis, Shleifer & Vishny, 1998; Black, 1986; De Long, Shleifer, Summers & Waldmann, 1990). The standard finance model's assumption that unemotional investors always force capital market prices to equal the rational present value of expected cash flows has thus been shown to be a poor fit to the stock market's historical patterns (Baker & Wurgler, 2007).

Investor sentiment – one of the main pillars of contemporary behavioural finance – can be broadly defined as a belief about future cash flows and investment risks that is not justified by the facts at hand (De Long et al., 1990). Many researchers also accept that investor sentiment is a key driver of asset prices (Hui, Zheng & Wang, 2013).

Theoretical models of this significant factor have been based on two important assumptions. First, two types of traders exist: noise traders, who have random beliefs, and rational arbitrageurs, who have Bayesian beliefs. Second, both types of traders are risk-averse, capital-constrained or otherwise impaired from freely buying and selling risky assets. Therefore, they represent a downward sloping demand for risky assets.

Both assumptions lead to an equilibrium through which noise traders' random beliefs can influence prices. As a result, rational investors – or arbitrageurs, as they are often called – are not as aggressive in forcing prices to respect fundamentals as standard models would suggest. In the language of contemporary finance, arbitrage has limits (Shleifer & Vishny, 1997). The literature shows a growing consensus that noise traders can induce large price movements and excessive volatility in the short run.

As a result, a growing number of empirical studies have sought to measure investor sentiment. Traditionally, empiricists have taken two approaches to measuring investor sentiment as most of these studies have identified direct and indirect sentiment measures (see Qiu and Welch [2006] for a literature review). While direct sentiment measures are derived from surveys asking individuals how they feel about stock market conditions and current or future economic conditions, indirect sentiment measures represent economic and financial variables that seek to capture investors' state of mind.

In recent years, innovative measures have been proposed that can handle the latest technological developments and consumers' social media usage patterns (Piñeiro-Chousa, López-Cabarcos & Pérez-Pico, 2016). According to Ho, Damien, Gu and Konana (2017, p. 69), the 'wisdom of the crowd provides market sentiments that can be a proxy for the market mood'. One measure, for example, relies on data on Internet search frequency by household members, as suggested by Beer, Herve and Zouaoui (2013); Da, Engelberg and Gao (2015) and Preis, Moat and Stanley (2013).

Recently, the literature on behavioural finance (Siganos, Vagenas-Nanos & Verwijmeren, 2017) has also recognised that the average sentiment level might contain important information. This level can be obtained based on neutral overall sentiment or a scenario in which half the population exhibits positive sentiment and the other half shows equally negative sentiment. Building on previous studies, the present study sought, therefore, to answer the following research question: Can positive and negative retail investor sentiment be directly measured through the Internet search behaviour of households – in this case, those in Portugal?

Previous studies have considered mainly negative measures of sentiment (Da et al., 2015; Tetlock, 2007) or combined measures of positive and negative sentiment (Mao, Counts & Bollen, 2015; Siganos et al., 2017). One exception is Uhl (2014), who highlighted the need to consider separately the influence of both positive and negative sentiment measures on stock activity. The cited author developed two indicators including positive and negative Reuters news stories and concluded that their impact on stock returns and volume is different. Thus, the second research question considered in the present study was: Do positive and negative sentiment both influence stock returns and volume?

This work extends previous research by proposing an innovative measure derived from a carefully honed list of relevant search terms. In a related study, Da et al. (2015) constructed the Financial and Economic Attitudes Revealed by Search (FEARS) index by aggregating a daily search volume index for keywords related to household financial and economic concerns – drawn from widely used finance terms dictionaries. The cited authors then tested how these search terms are used in practice. Beer et al. (2013), in turn, proposed a novel measure of French investor sentiment based on search volume data. Although these and other previous studies included a proxy for investor pessimism only, the present study included both positive and negative search terms.

We followed a top-down approach that made use of aggregate sentiment and its effect on market returns. Our research design combined the secondary research methodology developed by Da et al. (2015), starting with economic- and financial-related search terms, as well as a qualitative methodology based on personal interviews. In addition, we constructed validity and reliability tests of the resulting sentiment indexes.

In order to test the predictive validity of these sentiment measures, this study also examined the influence of both sentiment measures on stock market activity, namely, returns and volume. The research built on Siganos et al.'s (2017) work by considering both sentiment measures instead of considering divergence of sentiment alone.

The remainder of this paper is organised into the following sections. As many investor sentiment indicators have been developed, section two provides an overview of these measures and their role in behavioural finance, with a particular emphasis on Internet search data. Section three discusses the data and methodology used to derive our Google-based sentiment index. Section four reports the empirical results regarding the predictive capabilities of the proposed measures. The last section summarises the findings.

2. Overview of Investor Sentiment Measures

Investor sentiment can be defined as 'a belief about future cash flows and investment risks that is not justified by the facts at hand' (Baker & Wurgler, 2007, p. 129). The question of whether investor sentiment affects stock markets long been debated in behavioural finance. For instance, Barberis et al. (1998), Black (1986) and De Long et al.'s (1990) studies modelled the role of investor sentiment in financial markets. Grounded in the field of behavioural finance, the proposed models assume that investors are swayed by sentiment (De Long et al., 1990). In addition, betting against sentimental investors is costly and risky since arbitrage has limits (Shleifer & Vishny, 1997).

Previous studies have produced three main findings (Baker & Wuger, 2007; Da et al., 2015). First, investor sentiment explains stock returns. Second, sentiment has a larger influence on stocks whose valuations are more subjective and difficult to subject to arbitrage (e.g. stocks with higher beta, higher volatility and greater downside risk). Third, investor sentiment is subject to reversals so that, first, increases – or decreases –

in sentiment correspond to low – or high – returns and, second, in the days following transactions, this relationship reverses itself.

According to Baker and Wurgler (2007, p. 130), 'academic attention has moved from studying "whether investor sentiment affects stock prices" to assessing "how to measure investor sentiment and quantify its effects".' However, the question of what constitutes a good measure of investor sentiment still remains unanswered.

Various attempts have been made to quantify investor sentiment and evaluate the effectiveness of available measures to explain and predict stock market activity. The resulting indicators have been based on different methodologies, have used different data sets and sources, have targeted different retail investors and have been available for different time periods.

These previous studies have taken two main approaches to measuring investor sentiment (e.g. Brown & Cliff, 2004a, 2004b; Shiller, 2000): market-based (i.e. indirect) and survey-based (i.e. direct) measures. However, in recent years, significant improvements have been made in investor sentiment tracking techniques, and new and innovative measures have been suggested that extract indicators of public mood directly from social media content and from Internet search data. Both indirect and direct sentiment measures have advantages and shortcomings.

The first approach applies market-based measures. These include individual investors' orders executed (Kaniel, Saar & Titman, 2008), aggregate mutual fund flows of domestic equity funds (Beaumont, van Daele, Frijns, Lehnert & Muller, 2008) and shifts in investment allocations to risky assets by retail investors relative to those of institutional investors (Edelen, Marcus & Tehrnian, 2010). Backer and Wurgler (2007), in particular, derived a composite index of sentiment extracted from different market variables. These were trading volume, dividend premiums, closed-end fund discounts, the number of – and first-day returns on – initial public offerings and equity shares in new issues. All these measures have been widely used in the academic literature because they are easily constructed and derived from objective, observable financial data.

Market-based measures have the primary advantage of being readily available at a relatively high frequency. As they can be observed in real time, they reflect both the power of market participants and the strength of bullish or bearish approaches. However, using indirect measures of investor sentiment also produces difficulties in terms of validity. Their main weakness is the need to build up a theory relating them to sentiment, as well as how to interpret them. As these measures are endogenous to the market and economic activities, they also may not exclusively measure investor sentiment. The process of isolating one measure from the others can prove to be a difficult – if not impossible – task.

In the second main approach to measuring investor sentiment, this is assessed with survey-based indices based on consumer and investor polls. International examples of these measures are the University of Michigan Consumer Sentiment Index (Otoo, 1999), the Investors Intelligence survey (Brown & Cliff, 2004a; Lee, Jiang & Indro, 2002), the European Economic Sentiment Indicator (Vieira, 2011) and surveys from the Conference Board and the University of Michigan Survey Research Center (Lemmon & Portniaguina, 2006). Measures have also been created specifically for Portugal, such as the Economic Sentiment Indicator, developed under the Joint Harmonised European Union Programme of Business and Consumer Surveys.

These measures attempt to capture the mood of the market, having been created from surveys that directly measure the sentiment of market participants by asking them about their expectations of the market and using standardised questions. The surveys take into account individuals' psychological dimensions (e.g. optimism, pessimism and neutrality). However, as survey-based measures are based on self-reported information provided by consumers and investors, these measures are subject to measurement errors.

In addition, as survey-based measures require a representative panel of target populations, the measures have the disadvantage of being costly to use. Survey measures are also often published only on a monthly or quarterly basis. The highest frequency indicators are published on a weekly basis. In Portugal, the most frequent measures are reported monthly. Since most survey opinions are gradually submitted during a week, a month or a quarter, the results do not correspond to investor sentiment at a given point in time but instead to a mix of recent and previous opinions.

Some authors have also proposed the integration of market- and survey-based measures to assess the relationship between investor sentiment and stock market activity or compare the measures' relative efficacy in predicting future stock returns (e.g. Baker & Wurgler, 2007; Beer & Zouaoui, 2013; Brown & Cliff, 2004b; Feldman, 2010; Qiu & Welch, 2006). Another type of innovative and non-standard measure is based on an amalgam of opinions, thus referred to as 'meta-measures'. Some of these measures gather data from traditional media sources, such as newspapers' daily content (Tetlock, 2007) and newsletter writers (Clarke & Staman, 1998; Fisher & Statman, 2000). For example, Tetlock, Saar-Tsechansky and Macskassy (2008) collected negative phrases in financial media studies of individual firms' accounting earnings and stock returns.

Other studies have gathered direct public mood data from social media content, that is, user-generated content. This has included messages in Internet chatrooms focused on stocks (Antweiler & Frank, 2004), blogs, large-scale twitter feeds (Dergiades, Milas & Panagiotidis, 2015), Facebook activity (Siganos et al., 2017), messages driven by Yahoo services (Das & Chen, 2007; Kim & Kim, 2015), microblogs (i.e. Twitter data) (Bollen, Mao & Zeng, 2010; Ho et al., 2017; Mao et al., 2015) or Wikipedia usage patterns (Moat et al., 2013).

Online behaviour is thus assumed to be representative of trends in the general population. Mao et al. (2015) derived an online investor sentiment indicator based on Twitter updates and Google search queries. Twitter and Google bullishness was found to be positively correlated to investor sentiment and well-known investor sentiment surveys.

As social media content might reflect public sentiment in real time, this content has become a popular source for analyses of economic and financial topics (Dergiades et al., 2015). These measures appear, therefore, to be a promising approach to generating data that are more flexible in terms of high frequency, high degree of detail, low cost and unprecedented scale (Mao et al., 2015). However, these measures require various operations to be performed before data analysis, including data input or selection of text analysis packages. A main shortcoming thus is data availability, and some ethical issues can also arise (Poynter, 2010).

As recently suggested by Da et al. (2015), Internet searches may possess interesting signalling properties as they are generated through agents' spontaneous behaviour. Internet search-based measures' increasing popularity has been stimulated by the data availability offered by Google Trends, which provides the online search volume of any query term submitted to Google since 2004. Da et al. (2015) highlighted several advantages associated with the use of search-based sentiment measures compared with former alternatives. They are available at a higher (i.e. daily or weekly) frequency than survey-based measures, and they are more transparent than other social media driven measures are. That is, search-based measures directly gauge behaviour instead of asking about it.

Therefore, search-based data has the potential to reveal the underlying beliefs of populations directly, and the data are extremely useful in financial applications. Da et al. (2015) used the Internet search volume of several queries related to household concerns (e.g. ‘recession’, ‘unemployment’ and ‘bankruptcy’) to construct the FEARS index as a new measure of investor sentiment. The cited study’s results provide support for De Long et al.’s (1990) noise trader model. Da et al.’s (2015) index can predict daily realised volatilities of exchange-traded funds even after accounting for the effect of variables such as the VIX index, volume and turnover. The index also offers an alternative sentiment measure and predicts daily fund flows from equities to fixed income and mutual funds, which is consistent with a ‘flight to quality’ effect in turbulent times.

Preis et al. (2013) analysed changes in Google query volumes for search terms related to finance in order to identify patterns that may be interpreted as ‘early warning signs’ of stock market moves. A list of 98 search terms related to stock markets was obtained by means of the Google Sets service, a tool available up to 2011 that identified semantically related keywords. The results identify an increase in Google search volumes for keywords related to financial markets before stock markets fall.

Beer et al. (2013) proposed a novel measure of French investor sentiment based on the volume of Internet searches. The cited authors attempted to use a proxy for investor sentiment based on the first principal component of eight negative search terms (i.e. ‘bankruptcy’, ‘debtor’, ‘deficit’, ‘inflation’, ‘liquidation’, ‘poverty’, ‘recession’ and ‘crisis’). Beer et al. (2013) showed that their French sentiment indicator (i.e. pessimism) correlates with other measures of sentiment, and it is associated with outflows from equity funds and inflows to treasury bonds. The results also indicate that this sentiment index leads to short-term return reversals that are more pronounced for smaller firms than for larger firms, which is consistent with the predictions of noise trader models.

Mao et al.'s (2015) recent cross-country study (i.e. the United States [US], United Kingdom, Canada and China) included only two search terms in analyses – 'bull market' and 'bear market' – as the terms are rarely used other than in a financial context. Collectively, previous studies' results thus highlight the advantages of using Internet search data frequency as a new approach to measuring retail investment sentiment. For example, Beer et al. (2013) and Da et al.'s (2015) studies used investor pessimism as a proxy for investor sentiment, applying negative economic search terms.

However, more recent studies in the field of behavioural finance have highlighted that combined positive and negative social media measures of sentiment can contain important information (Mao et al., 2015; Siganos et al., 2017) and influence stock market activity. Furthermore, the literature reviewed for this study revealed that room exists for the development of both positive and negative sentiment indicators based on Google Search data.

3. METHODOLOGY

3.1 Google search volume data

Following Da et al.'s (2015) lead, the present research assumed that household attitudes, as represented by Internet search behaviour, can be considered a measure of sentiment. Market-level sentiment was, therefore, directly measured through weekly Internet search behaviour of households. By aggregating the volume of queries associated with household concerns, we sought to develop a Google-based measure of consumer sentiment.

Google is the most popular search engine in the world. This engine provides the Google Search Volume Index (GSVI) of search terms through its Google Trends product (see http://www.google.com/trends). When users input search terms into Google Trends, the application returns the search volume history for that term, scaled by the time-series maximum (i.e. scalar).

The key to the construction of a Google-based sentiment index is the identification of an appropriate list of sentiment-revealing search terms. Therefore, we followed Da et al.'s (2015) proposed method of building a list of search terms that reveal sentiment towards economic conditions. A set of 149 primitive terms classified as economic words with either positive or negative sentiment were identified from widely used dictionaries in the literature on finance and textual analysis (i.e. *Harvard IV-4 Dictionary* and *Lasswell Value Dictionary* (Tetlock, 2007; Tetlock et al., 2008). This list was compared with the final list of 118 terms obtained by Da et al. (2015), who identified the top 10 search terms associated with each one of 149 primitive words and eliminated non-economic terms and terms with too few valid GSVIs. The final list included both positive (e.g. 'entrepreneurship') and negative (e.g. 'crisis') terms.

Google Trends allows users to restrict GSVI results to specific countries. The GSVI results were, therefore, restricted to Portugal for the present study as the dependent variables of interest were related to the Portuguese stock market and the measures we sought to develop represented the sentiment of Portuguese households.

In terms of periodicity, Google Trends offers two options when downloading GSVI data. The standard mode allows users to download data on a weekly basis, in which the data are scaled by the weeks in January 2004. Google Trends also provides GSVIs on a daily basis when users download data for a time window less than or equal to a quarter of a year. In this case, the daily GSVIs in a particular quarter are scaled by the time series' maximum GSVI for that quarter. Although Da et al. (2015) used daily data, in the present study, we downloaded weekly GSVIs, as most search terms were not available for shorter periods.

3.2 Qualitative study

As Da et al.'s (2015) study was performed in the US, personal interviews were conducted to account for possible cross-country differences. The present research also used this approach to add content validity to the final list of search terms. The target respondents in this qualitative study were students and professionals in the fields of economics and management, living in Portugal at the time of the survey. The respondents were divided into three age groups: 18 to 24 years old (number [N] = 30), 25 to 44 years old (N = 30) and 45 or more years old (N = 30). Fifty-two respondents are male (58%), and 38 are female (42%).

The interview guide included three questions. The respondents were asked to list which search terms they would put into an Internet search engine if they wanted to get information 'about the state of the European or world economy', 'the state of the Portuguese economy' and 'the Portuguese financial markets'. In order to summarise the results, a word cloud[1] was generated for each question (see Exhibits 1–3). The most frequent words for the first question were World Bank, European Economy, World Economy (N=36, 40%), Economic Growth (N=33, 37%) and IMF (N=30, 33%). For the second question, the most frequent words were Portuguese Economy (N=60, 67%), Salaries, State of Portuguese Economy (N=36, 40%), Crises in Portugal, IRS levels, Unemployment (N=30; 33%). Finally, for the third question the most frequent words

[1] The clouds give greater prominence to search terms that appear more frequently in the source text.

were: Financial Markets, Lisbon Stock Exchange (N=45, 50%), Prices, PSI 20 (N=42, 47%), Bank of Portugal, Euribor, Euronext, Interest Rate (N=33, 37%).

Insert Exhibit 1 here.

Insert Exhibit 2 here.

Insert Exhibit 3 here.

After removing duplicated terms, this two-step approach resulted in 345 search terms. In order to evaluate how these economic words are used in practice, we also considered up to 10 related terms associated with these 345 terms. This procedure left us with 979 final search terms. A final list of 105 search terms resulted from the removal of terms that produced insufficient data and that were not clearly related to economics or finance.

The GSVIs for each of these 105 search terms were downloaded from Google Trends from January 2009 to February 2014[2] and restricted to Portugal. Exhibit 4 shows the GSVI for the search terms 'crisis' and 'unemployment' for the period under analysis.

Insert Exhibit 4 here.

3.3 Data transformations

Because it was simpler to interpret, we used the logarithm of GSVI, denoted by SVI for each search term *j* for week *t*, as shown in Equation (1):

$$SVI_{j,t} = ln(GSVI_{j,t}) \tag{1}$$

[2] This study used historical data from January 2009 to February 2014 in order to include a time frame after the subprime crisis.

Next, the weekly change in search term j was defined as shown in Equation (2):

$$\Delta SVI_{j,t} = SVI_{j,t} - SVI_{j,t-1} \quad (2)$$

To make the final list of terms comparable and account for outliers, seasonality and heteroscedasticity in the data, several further transformations were performed. First, to mitigate our concerns about outliers, we winsorised each series at the 5% level (i.e. 2.5% in each tail). Then, we tested for the presence of intra-year seasonality by performing 105 one-way analysis of variance (ANOVA) tests.[3]

The null hypothesis of equality of means across the 12 months was rejected only for 20 search terms.[4] For these terms, we regressed $\Delta SVI_{j,t}$ on monthly dummies and kept the residuals to obtain deseasonalised weekly changes in search volumes. Finally, to address any heteroscedasticity in the data and to make the time series comparable, we standardised each time series by the time-series standard deviation. This procedure allowed us to obtain an adjusted daily change in search volume for each of the 105 search terms, denoted by $\Delta ASVI_{j,t}$.

3.4 Google-based sentiment index

The final step in the construction of the Google-based sentiment index was identifying search terms that are the most important for returns. We determined the historical relationship between each term and contemporaneous market returns for all 105 series of adjusted search volume terms.[5] We then selected the 20 search terms with

[3] Out of the large number of methods available to account for data seasonality, we followed the two-step approach (i.e. ANOVA and multiple regressions with dummy variables) proposed by Da et al. (2015), as this is appropriate when dealing with Google-based data.
[4] The search terms were: 'European Commission', 'accounts', 'crisis', 'deficit', 'state budget', 'dividends', 'economy', 'enterprises', 'finance', 'INE', 'IRS', 'IRC', 'bonds', 'OECD', 'political parties', 'passive', 'poverty', 'prices', 'salaries' and 'consume'.
[5] Market returns were defined as the weekly changes in closing prices of the PSI-20 TR Index.

the highest correlation with the market as those having a *t*-statistic higher than 2.5 (10) and those having a *t*-statistic less than -2.5 (10).

The final list of positively correlated terms ($\Delta ASVI^{+}$) included ‘stocks’, ‘consume’, ‘credit’, ‘GDP’, ‘Lisbon stock market’, ‘dividends’, ‘profits’, ‘investment’, ‘entrepreneurship’ and ‘partnership’. Exploratory factor analysis using a principal components method revealed the existence of one factor that explains 59% of the variance. The Cronbach’s alpha coefficient for these 10 items is 0.79, thereby meeting the cut-off point of 0.70 for exploratory studies. Therefore, the derived items satisfy the reliability criteria.

The Google-Based Positive Sentiment Index (PSI) was defined as shown in Equation (3):

$$PSI_t = \sum_{j=1}^{10} \Delta \mathrm{ASVI}_{j,t}^{+}/10 \tag{3}$$

The final list of negatively correlated terms ($\Delta ASVI^{-}$) comprised ‘austerity’, ‘taxes’, ‘rents,’ ‘Euribor’, ‘crisis’, ‘debt’, ‘finance’, ‘gold price’, ‘unemployment’ and ‘poverty’. Exploratory factor analysis using a principal component method revealed the existence of one factor that explains 62% of the variance. The Cronbach’s alpha coefficient for these 10 items is 0.84.

The Google-based Negative Sentiment Index (NSI) was defined as shown in Equation (4):

$$NSI_t = \sum_{j=1}^{10} \Delta \mathrm{ASVI}_{j,t}^{-}/10 \tag{4}$$

During 2010 and the first half of 2011, the search for queries that match the PSI was stronger than the search for queries that fall within the NSI.[6] Over the second half of 2011 and during 2012 and the first half of 2013, the search for both queries exhibited the opposite behaviour.

The search volume for the list of negative terms (i.e. 52-week moving average) was more strongly correlated with the aforementioned Economic Sentiment Indicator (r = -0.74) than the search volume for the list of positive search terms (r = 0.68). The search volume for the PSI and NSI's search terms thus revealed information shared with the Economic Sentiment Indicator. However, Google-based indicators offer additional advantages as they are available at a greater frequency and lower cost (see Exhibit 5).

Insert Exhibit 5 here.

4. RESULTS

4.1 Google-based sentiment index and stock market activity

This subsection presents an analysis of the results for the relationship between the proposed Google-based sentiment index and Portuguese stock market activity – based on weekly data. In this analysis, we used a top-down approach that made use of aggregate sentiment and its effect on market returns.

Granger causality tests revealed that only two null hypotheses need to be rejected, namely, the PSI does not Granger cause the PSI-20 TR returns and the NSI does not Granger cause the PSI-20 TR returns. We focused, therefore, on how search

[6] In order to facilitate the comparison of the search volume data for the two groups of variables, data were obtained from two groupings simultaneously.

activity influences stock market returns. The empirical studies reviewed also support this modelling approach (see Exhibit 6).

Insert Exhibit 6 here.

We modelled the PSI-20 TR[7] weekly returns as a function of the Google-based sentiment indexes. The independent variables were the PSI and NSI. The control variables included one-week lagged returns and realised volatility. Returns, denoted by RET_t, were defined as shown in Equation (5):

$$RET_t = ln\left(\frac{P_t}{P_{t-1}}\right), \tag{5}$$

in which P_t was the observed weekly closing price of the PSI-20 TR Index.

We based our analysis on previous studies of stock market volatility (Vlastakis & Markellos, 2012) and proxy volatility using the standard deviation of returns. The realised volatility at week *t* (RV_t) was computed from daily data, in which RET_j^2 corresponded to the squared return of the *i*th stock for day *j*, as shown in Equation (6):

$$RV_t = \sqrt{\sum_{j=1}^{5} RET_j^2} \tag{6}$$

Our results suggest a positive (negative) contemporaneous relationship between the PSI (NSI) and the PSI-20 TR Index returns (see Table 1). The null hypothesis of equality between the absolute value of the PSI and NSI's coefficients was rejected (F = 10.03; *p* = 0.00). This indicates that the impact of the PSI is stronger than the impact of the NSI.

We based the final selection of the terms that comprise the Google-based search index on the terms' correlations with market returns. Thus, in a second step, we further

---

[7] We also estimated all the models for the PSI-20 Index. As the results are quite similar, they are not presented in this paper.

estimated two models in order to test if the obtained Google-based sentiment indexes relate to stock market volume and abnormal volume. Stock market volume (VOL_t) was measured by the weekly log change in turnover, as shown in Equation (7):

$$VOL_t = ln\left(\frac{Vol_t}{Vol_{t-1}}\right) \quad (7)$$

Abnormal volume ($\boldsymbol{Abn\_VOL_t}$), was defined as suggested by Barber and Odean (2008), which is shown in Equation (8):

$$Abn\_VOL_t = \frac{lnVol_t}{\sum_{j=1}^{52} lnVol_{t-j}} \quad (8)$$

Both positive and negative versions of our Google-based sentiment indexes are statistically significant (see Exhibit 7). The impact of the PSI on market volume and abnormal volume is positive, and the impact of the NSI is negative.[8] The null hypotheses of equality between the absolute coefficients of the PSI and NSI were rejected in both cases.

Insert Exhibit 7 here.

4.2 Predictability of stock market activity

The empirical results presented in the previous subsection support the conclusion that the proposed Google-based sentiment indexes exhibit a contemporaneous association with stock market activity, namely, returns, volume and abnormal volume. We next tested whether the PSI and NSI predict these stock market variables. In these model formulations, all control variables were lagged one week.

[8] However, the Google-based sentiment measures failed to explain the abnormal return for the PSI-20 Index, which was estimated by means of a market model and quantified using the residuals of the market model, in which market-wide movement was proxied by the STOXX Europe Index returns.

The results shown in Exhibits 8 and 9 reveal that the one-week lagged NSI is significant (i.e. at the 5% level) for the returns model. However, the one-week lagged PSI is not significant for these models. The results for the volume and abnormal volume regressions reveal that only the one-week lagged PSI is significant in both models.

Insert Exhibit 8 here.

Insert Exhibit 9 here.

5. Conclusion and Discussion

This study sought to answer two research questions. The first was: Can retail investor sentiment be directly measured through Internet search behaviour, in this case, of Portuguese households? Following Da et al.'s (2009, 2015) example, the present research assumed that household attitudes – as revealed by Internet search behaviour – can be considered a measure of sentiment. We, therefore, employed a two-step approach combining the final list of search terms created by Da et al. (2009), a list of economic and finance terms in widely used dictionaries and the results of 90 personal interviews with management students and professionals living in Portugal. Thus, the final list of terms presented previously has content validity regarding the measurement of retail investor sentiment.

Using historical search data from the period between January 2009 and February 2014, this research identified two sets of search terms that are positively and negatively correlated with PSI-20 TR returns. Factor analysis confirmed that Google search sentiment is a multidimensional construct, comprising both positive and negative components. The reliability of each component was also confirmed by its Cronbach's alpha coefficient.

The volume of positive and negative sentiment search terms revealed information shared with the previously-developed Economic Sentiment Indicator. However, Google-based indicators offer further advantages as compared to direct measures of economic sentiment since the former are more frequently available at a lower cost.

Tetlock (2007) sought to develop different types of media sentiment measures: pessimism, negative (outlook) and weakness. However, the cited author argues that negative search terms in English are the most useful for identifying sentiment. Beer et al. (2013) and Da et al. (2009, 2015) similarly derived their Google-based sentiment indexes based solely on negative queries. In contrast, Uhl (2014) developed measures of positive and negative media sentiment. Siganos et al. (2017), in turn, introduced the concept of divergence of sentiment to the literature on behavioural finance and encouraged the use of this concept instead of average sentiment levels.

The present study concluded that retail investor sentiment can be directly measured through the Internet search behaviour of Portuguese households. In addition, this research produced innovative insights using two sentiment measures based on Google search data: the PSI and NSI.

The second research question this study addressed was: Do positive and negative sentiment both influence stock returns and volume? The results reveal that an increase in positive (negative) sentiment is associated with a contemporaneous increase (decrease) in stock returns. This finding is unexpected as the search terms that comprise the sentiment indexes were selected based on their historical correlation with stock returns, as was done in Da et al.'s (2015) research. The cited authors' results were estimated daily and only considered negative search terms. Mao et al. (2015) also concluded a correlation exists between Google search's bullish indicator and weekly stock market returns.

In addition, the one-week lagged NSI predicts market returns. This finding is in accordance with Da et al.'s (2009, 2015) study, which found that only negative search terms predict stock market returns.

Regarding trading volume and abnormal trading value, the present proposed model confirmed that the one-week lagged PSI is statistically significant. De Long et al. (1990) and Tetlock (2007) argue that, if media pessimism predicts future investor sentiment, unusually high or low levels of pessimism forecast a high trading volume. The current study thus found evidence that high or low levels of positive sentiment predict one-week lagged returns.

This research's findings have managerial implications. The results confirm the validity of developing measures of sentiment based on Internet searches, which encompass both positive and negative sentiments. These sentiment measures are significantly useful for the short-term prediction of market returns and market volume.

The results correspond to a new stream of research in the literature on sentiment that encourages the use of Internet search queries. The main appeal of using Internet searches is that they are generated through the spontaneous behaviour of households, thereby offering interesting signalling properties. The data are updated weekly, they are free and unrestricted and they can be collected in real time, as well as offering the possibility of large samples. In addition, the proposed combination of big data on search query frequency with financial trading data can offer new insights that help develop a better understanding of the complex collective behaviour of households, in this case, in Portugal.

Some limitations of the present study and the resulting need for additional research need to be considered. The findings confirm that the proposed Portuguese PSI and NSI offer a good explanation of market sentiment specifically for the study period in question. Moreover, these Portuguese sentiment indicators correlate closely with well-known alternative sentiment measures.

However, to validate these findings further, the stability of the search terms obtained with data from 2009 to 2014 could require further testing. In addition, as the final list of search terms diverged from those proposed by Beer et al. (2013) and Da et al. (2015), future studies need to confirm the results' cross-country validity. Researchers should test whether similar indexes can be defined that can be implemented in international comparisons. Further cross-country analysis also needs to be done because investor sentiment in theory can be based on not only household concerns but also positive economic search terms.

6. References

Antweiler, W., & Frank, M. (2004). Is all that talk just noise? The information content of Internet stock message boards. Journal of Finance, 59(3), 1259–1294.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. Journal of Economic Perspectives, 21(2), 129–151.

Barber, B., & Odean, T. (2008). All that glitters: The effect of attention on the buying behavior of individual and institutional investors. Review of Financial Studies, 21(2), 785–818.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. Journal of Financial Economics, 49(3), 307–343.

Beaumont, R., van Daele, M., Frijns, B., Lehnert, T., & Muller, A. (2008). Investor sentiment, mutual fund flows and its impact on returns and volatility. Managerial Finance, 34(11), 772–785.

Beer, F., Herve, F., & Zouaoui, M. (2013). Is Big Brother watching us? Google, investor sentiment and the stock market. Economics Bulletin, 33(1), 454–466.

Beer, F., & Zouaoui, M. (2013). Measuring stock market investor sentiment. The Journal of Applied Business Research, 29(1), 51–68.

Black, F. (1986). Noise. Journal of Finance, 41, 529–543.

Bollen, J., Mao, H., & Zeng, X.-J. (2010). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1–8.

Brown, G., & Cliff, M. (2004a). Investor sentiment and the near-term stock market. Journal of Empirical Finance, 11(1), 1–27.

Brown, G., & Cliff, M. (2004b). Investor sentiment and asset valuation. Journal of Business, 78 (2), 405–440.

Da, Z., Engelberg, J., & Gao, P. (2009). In search of attention. The Journal of Finance, 66(5), 1461–1499.

Da, Z., Engelberg, J., & Gao, P. (2015). The sum of all FEARS: Investor sentiment and asset prices. The Review of Financial Studies, 28(1), 1–32.

Das, S., & Chen, M. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the Web. Management Science, 53(9), 1375–1388.

De Long, J., Shleifer, A., Summers, L., & Waldmann, R. (1990). Noise trader risk in financial markets. Journal of Political Economy, 98(4), 703–738.

Dergiades, T., Milas, C. & Panagiotidis, T. (2015). Tweets, Google trends, and sovereign spreads in the GIIPS. Oxford Economic Papers, 67(2), 406–432

Edelen, R. M., Marcus, A. J., & Tehrnian, H. (2010). Relative sentiment and stock returns. Financial Analysts Journal, 66(4), 20–32.

Feldman, T. (2010). A more predictive index of market sentiment. The Journal of Behavioral Finance, 11(4), 211–223.

Ho, C.-S., Damien, P., Gu, B., & Konana, P. (2017). The time-varying nature of social media sentiments in modeling stock returns. Decision Support Systems, 101(3), 69–81.

Hui, E., Zheng, X., & Wang, H. (2013). Investor sentiment and risk appetite of real estate security market. Applied Economics, 45(19), 2801–2807.

Kaniel, R., Saar, G., & Titman, S. (2008). Individual investor trading and stock returns. The Journal of Finance, 63 (1), 273–310.

Keynes, J. (1936). The general theory of employment, interest and money. Cambridge, UK: Macmillan’s Cambridge University Press.

Kim, S.-H., & Kim, D. (2015). Investor sentiment from internet message postings and the predictability of stock returns. Journal of Economic Behavior & Organization, 107, 708–729.

Kräussl, R., & Mirgorodskaya, E. (2017). Media, sentiment and market performance in the long run. The European Journal of Finance, 23(11), 1059–1082.

Lee, W., Jiang, C., & Indro, D. (2002). Stock market volatility, excess returns, and the role of investor sentiments. Journal of Banking and Finance, 26(12), 2277–2299.

Lemmon, M., & Portniaguina, E. (2006). Consumer confidence and asset prices: Some empirical evidence. Review of Financial Studies, 19(4), 1499–1529.

Mao, H., Counts, S., & Bollen, J. (2015). Quantifying the effects of online bullishness on international financial markets, statistics papers, No. 9/July 2015. Frankfurt: European Central Bank.

Moat, H. S., Curme, C., Avakian, A., Kenett, D. Y., Standley, H. E., & Preis, T. (2013). Quantifying stock market moves. Scientific Reports, 3, 1801.

Otoo, M. W. (1999). Consumer sentiment and the stock market, No 1999-60, finance and economics discussion series. Washington, DC: Board of Governors of the Federal Reserve System.

Piñeiro-Chousa, J. R., López-Cabarcos, M. A., & Pérez-Pico, A. M. (2016). Examining the influence of stock market variables on microblogging sentiment. Journal of Business Research, 69(6), 2087–2092.

Poynter, R. (2010). The handbook of online and social media research: Tools and techniques for market researchers. Hoboken, NJ: John Wiley & Sons.

Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying trading behavior in financial markets using Google Trends. Scientific Reports, 3(1684). Retrieved from https://www.nature.com/articles/srep01684 (accessed 15 May 2017).

Qiu, L., & Welch, I. (2006). Investor sentiment measures, working paper. Providence, RI: Brown University.

Shiller, R. J. (2000). Measuring bubble expectations and investor confidence. The Journal of Psychology and Financial Markets, 1(1), 49–60.

Shleifer, A., & Vishny, R. (1997). The limits of arbitrage. Journal of Finance, 52, 35–55.

Siganos, A., Vagenas-Nanos, E., & Verwijmeren, P. (2017). Divergence of sentiment and stock market trading. Journal of Banking and Finance, 78, 130–141.

Tetlock, P. (2007). Giving content to investor sentiment: The role of media in the stock market. Journal of Finance, 62(3), 1139–1168.

Tetlock, P., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. Journal of Finance, 63(3), 1437–1467.

Uhl, M. W. (2014). Reuters sentiment and stock returns. The Journal of Behavioral Finance, 15(4), 187–298.

Vieira, E. S. (2011). Investor sentiment and the market reaction to dividend news: European evidence. Managerial Finance, 37 (12), 1213–1245.

Vlastakis, N., & Markellos, R. (2012). Information search and stock market volatility. Journal of Banking and Finance, 36(6), 1808–1821.

Exhibit 1. Word cloud of search terms about the state of the European and world economies (N = 90)

Source: Authors

Exhibit 2. Word cloud of search terms about the state of Portugal’s economy (N = 90)

Source: Authors

Exhibit 3. Word cloud of search terms about the state of financial markets (N = 90)

Source: Authors

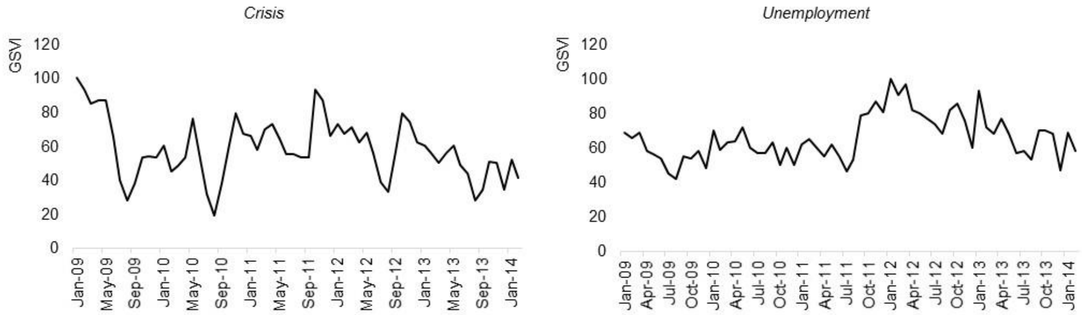

Exhibit 4. Google search volume monthly data from January 2009 to February 2014

Source: Google Trends (n.d.)

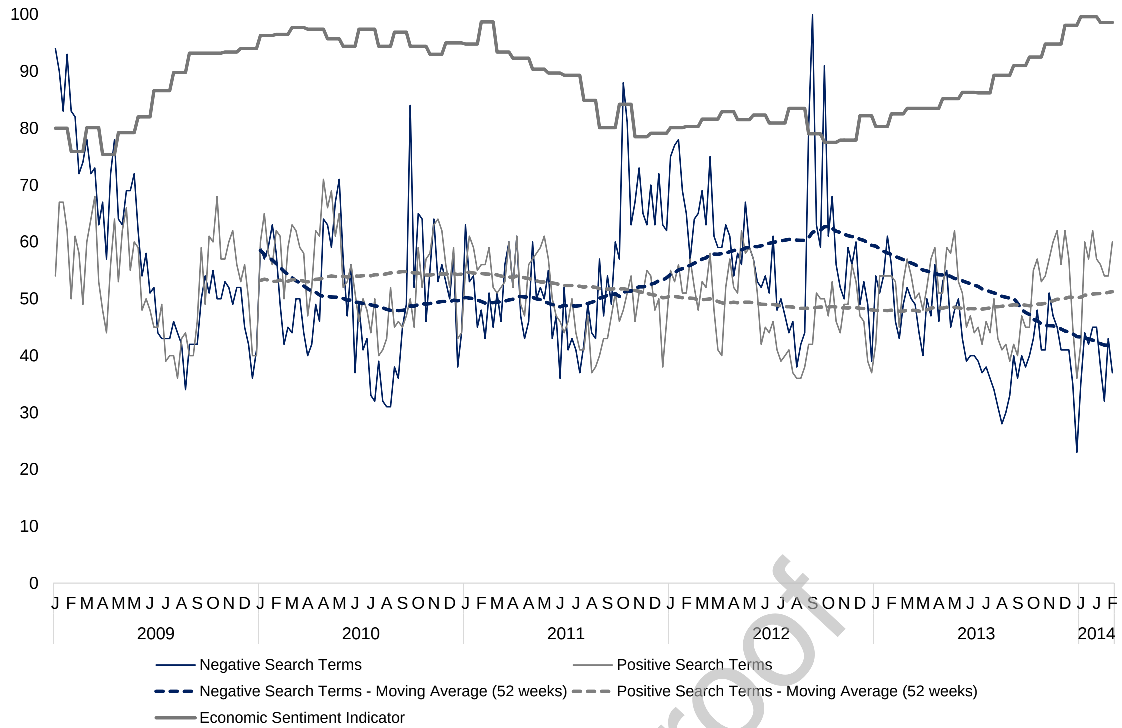

Exhibit 5. Positive search terms, negative search terms and the Economic Sentiment Indicator

Source: Adapted from Google Trends (n.d.); Banco de Portugal (n.d.) (authors' calculations)

Exhibit 6. Granger causality tests

Null Hypothesis	F-Statistic	P-Value
Negative sentiment does not Granger cause PSI 20 TR returns.	5.75	0.00
PSI-20 TR returns do not Granger cause negative sentiment.	1.09	0.34
Positive sentiment does not Granger cause PSI 20 TR returns.	4.25	0.00
PSI 20 TR returns do not Granger cause positive sentiment.	0.75	0.47

Source: Authors

Exhibit 7. Positive and negative sentiment, PSI-20 TR returns, volume and abnormal volume

Model	***RET_t***		**VOL_t**		**Abn_VOL_t**	
	β	**SE**	**β**	**SE**	**β**	**SE**
Constant	0.006	0.004*	0.000	0.002	0.341	0.064**
PSI_t	0.064	0.010***	0.066	0.013***	0.49	0.13***
NSI_t	-0.048	0.010***	-0.045	0.014***	-0.18	0.103*
RET_{t-1}	-0.010	0.005**	–	–	–	–
VOL_{t-1}	–	–	-0.011	0.054	–	–
Abn_VOL_{t-1}	–	–	–	–	0.417	0.548
RV_t	0.003	0.002*	0.009	0.005*	0.741	0.049**
Adj. R-squared	0.150		0.138		0.492	

Notes. *, ** and *** = statistically significant at the 1%, 5% and 10% levels; Newey-West HAC standard errors and covariance were used in the estimation; RET = return; VOL = volume; Abn_VOL = abnormal volume; RV = realised volatility.

Source: Authors

Exhibit 8. One-week PSI and NSI, PSI-20 TR returns, volume and abnormal volume

Model	***RET_t***		***VOLt***		***Abn_VOLt***	
	β	**SE**	**β**	**SE**	**β**	**SE**
C	0.003	0.003	0.099	0.040**	0.414	0.058
PSI_{t-1}	-0.014	0.012	0.565	0.132***	0.417	0.129***
NSI_{t-1}	-0.043	0.01***	-0.175	0.106*	0.153	0.103
RET_{t-1}	-0.011	0.006**	–	–	–	–
VOL_{t-1}	–	–	0.245	0.055***	–	–
Abn_VOL_{t-1}	–	–	–	–	0.665	0.045***
RV_{t-1}	0.009	0.006	0.050	0.010***	-0.001	0.001*
Adj. R-squared	0.060		0.174		0.483	

Notes. *, ** and *** = statistically significant at the 1%, 5% and 10% levels; Newey-West HAC standard errors and covariance were used in the estimation; RET = return; VOL = volume; Abn_VOL = abnormal volume; RV = realised volatility.

Source: Authors

Exhibit 9. PSI and NSI, PSI-20 TR returns, volume and abnormal volume

Model	RET_t		VOL_t		Abn_VOL_t	
	β	**SE**	**β**	**SE**	**β**	**SE**
Constant	0.006	0.004*	-0.028	0.040	3.828	0.747**
PSI_t	0.063	0.014***	0.089	0.014***	0.468	0.13***
NSI_t	-0.046	0.013***	-0.273	0.108***	-0.190	0.102*
PSI_{t-1}	-0.005	0.001	0.647	0.133***	0.481	0.122***
NSI_{t-1}	-0.010	0.005**	0.138	0.108	0.1617	0.102
RET_{t-1}	-0.019	0.009**	–	–	–	–
VOL_{t-1}	–	–	-0.011	0.054	–	–
Abn_VOL_{t-1}	–	–	–	–	0.714	0.057***
RV_{t-1}	-0.003	0.002	–	–	-0.004	0.002*
Adj. R-squared	0.157		0.154		0.527	

Notes. *, ** and *** = statistically significant at the 1%, 5% and 10% levels; Newey-West HAC standard errors and covariance were used in the estimation; RET = return; VOL = volume; Abn_VOL = abnormal volume; RV = realised volatility.

Source: Authors