

**CAN GOOGLE TRENDS AFFECT SENTIMENT
OF INDIVIDUAL INVESTORS?
THE CASE OF THE UNITED STATES**

Aleksandra Rutkowska, Agata Kliber

Abstract. As the empirical studies show, investor sentiment is a significant factor in financial markets. The large-scale development of the technology has led to widespread access to information in real time (also to individual investors), which in turn has also led to the inflow of Big Data to market analysis. One of the sources of such data is the ability to track the phrases searched for in the web search engines. In our research we verify whether investor sentiment is affected by, among others, a daily Google keyword search called “Google Trends”. We consider measures of US investors’ sentiment calculated from survey studies – the AAII index. We investigate changes of sentiment and its volatility, which can be interpreted as nervousness of the market participants. We estimate a set of GARCH models with explanatory variables in conditional mean and variance. We confirm that negative keyword searches are connected with the decline of the investor confidence. The overall effect of a negative search is stronger than positive. Older searches have a weaker influence on investor sentiment than new ones – no lagged search proved to be significant.

Keywords: investor sentiment, GARCH models, behavioural finance, Google Trends.

JEL Classification: G40, G11, G17

DOI: 10.15611/me.2018.14.05

1. Introduction

The empirical studies have proved that investor sentiment is a significant factor in financial markets. The individual investor has widespread access to information almost in real time thanks to the large-scale technological development, which in turn has also led to the inflow of Big Data into market analysis. One of the sources of such data is the ability to track the phrases searched for in the web search engines.

The main goal of our study was to check the reflection of Google search in AAII sentiment based on selected keywords with a clearly defined sentiment. We examine whether behavioural aspects, like risk aversion, can be

Aleksandra Rutkowska

Poznan University of Economics and Business
e-mail: aleksandra.rutkowska@ue.poznan.pl
ORCID: 0000-0002-2111-7764

Agata Kliber

Poznan University of Economics and Business
e-mail: agata.kliber@ue.poznan.pl
ORCID: 0000-0003-1996-5550

identified through the search patterns. The following research hypotheses were formulated:

H1: Negative keyword search is connected with a decline of investor confidence, while positive – with its increase (effect in conditional mean).

H2: The overall effect of a negative search should be stronger (risk aversion, leverage effect).

H3: Older searches should have a weaker influence on investor sentiment than the newer ones (according to the law of small numbers).

We studied these hypotheses using a set of GARCH models with explanatory variables in conditional mean and variance. As variables we chose frequency of search phrases in the Google search engine available through the Google Trends tool (further: GT) and survey-based sentiment measures published by the American Association of Individual Investors (further: AAII).

2. Investor sentiment

According to Modern Portfolio Theory, financial markets are efficient and investors make rational decisions based on the available information. Moreover, everyone has the same access to information so all securities' prices reflect all the available information. However, several empirical studies reject the hypothesis of an efficient market. For instance, studies in behavioral finance have shown the effect of irrational investors on the market. Tversky and Kahneman as early as 1974 [Tversky, Kahneman 1975] found three kinds of bias in judgment and decision-making which concern assessing probabilities and predicting values: representativeness, availability and adjustment by anchoring.

2.1. Sentiment effects

De Long et. al [De Long et al. 1990] argue that some market anomalies can be explained by the idea of the so-called noise trader. As a result, prices can diverge significantly from fundamental values and noise traders may earn a higher expected return than rational investors. Moreover, Brown and Cliff [Brown, Cliff 2004] showed that the strongest relationships exist between measures of institutional sentiment and large stocks. The next study by the same researchers [Brown, Cliff 2005] showed that overly optimistic (pessimistic) investors drive prices above (below) fundamental values. Market pricing errors implied by an independent valuation model are positively related to sentiment. Future returns over multi-year horizons are negatively related to senti-

ment. Baker and Wurgler [Baker, Wurgler 2006] confirmed this result, showing that the cross-section of future stock returns is conditional upon beginning-of-period proxies for sentiment. The results are more noticeable in the case of stocks of companies that are hard to price and thus difficult to arbitrage. The sentiment effect is present not only in the US market. In [Schmeling 2009], 18 countries around the globe were investigated and the authors find the significant impact of investor sentiment on aggregate stock returns across countries. Moreover, using cross-sectional regressions, Schmeling provided the first evidence that the impact of sentiment on stock returns was stronger in countries of less integrated markets and in countries that (due to cultural aspects) were more prone to herd-behavior. Hibbert et al. [Hibbert et al. 2008] showed empirically that the short-term dynamic between the S&P 500 index return and the changes in implied volatility, both at daily and intra-daily frequency, can be explained by the representativeness effect and extrapolation bias concepts of behavioural finance.

2.2. Sentiment measures

There are several direct and indirect sentiment measures used in the above-mentioned studies. Among the **direct measures** – the most commonly used are two surveys: a survey conducted by the American Association of Individual Investors (AAII) and also the Investors Intelligence data as a proxy for institutional sentiment. The AAII polls its members every week (between 26 and 422 respondents) and has done so since July 1987. When we refer to the Investors Intelligence sentiment index, we are generally talking about advisor sentiment available since 1965, which is calculated through categorizing approximately 150 market newsletters. These direct measures were used, among others, in the following research: [Brown, Cliff 2004; Fisher, Statman 2000].

Indirect proxies of sentiment are based on recent market performance, type of trading activity or derivatives trading activity. Baker and Wurgler [Baker, Wurgler 2006] formed a composite index of sentiment that is based on the common variation in six underlying proxies for sentiment: the closed-end fund discount, NYSE share turnover, the number and average first-day returns on Initial Public Offerings (further: IPOs), the equity share in new issues, and the dividend premium. Among the most popular indirect sentiment measures the following can be also mentioned:

- *bullish measure*: the percentage change in margin borrowing, the number of new highs to new lows, technical indicators like the ratio of the number of advancing issues to declining issues, the ratio of short sales to total sales;

- *bearish measure*: the percentage change in short interest, the ratio of odd-lot sales to purchases, the ratio of CBOE equity put to call trading volume.

Brown and Cliff [Brown, Cliff 2004] checked also if some proxies were significantly related to the survey data. However the study results showed that in some cases the opposite relationships hold, for example sentiment is associated with increases in short interest and relatively high specialist short selling. Odd-lot selling is positively related to sentiment, in conflict with the conventional interpretation. The closed-end fund discount (measured here as a premium to net asset value) is significantly and negatively related to survey sentiment, IPOs tend to occur during bullish periods. The mutual fund data suggest that during times of high sentiment, investors are putting money into mutual funds and the funds are holding relatively little cash. In recent years there has been an increase in the role of the Internet, so the search for proxy of sentiment has been expanded by data available from the new media.

2.3. The role of the Internet

Antweiler and Frank [Antweiler, Frank 2004] analysed text messages (approximately 1.5 million) from Yahoo! Finance and RagingBull.com on 45 large firms in the calendar year 2000. The authors find that a positive shock to message board posting predicts negative company returns on the next trading day and that investor sentiment from Internet posting messages has a predictive power for volatility and trading volume. The same data source, but extended to more than 32 million messages on 91 firms over the period January 2005 to December 2010 were used in [Kim, Kim 2014]. The analysis showed that posting¹ is positively affected by prior stock price performance, but there was no significant effect in the other direction.

Another social-media data source for economic analyses is Twitter. Bollen's et. al [Bollen et. al 2011] results indicate that the accuracy of DJIA predictions can be significantly improved by the inclusion of specific public mood dimensions based on analyses of the text content of tweets. For instance, Meinusch and Tillmann [Meinusch, Tillmann 2017] used Twitter data to quantify user beliefs about the timing of the exit from Quantitative Easing (tapering) and to assess the impact of the beliefs on asset prices. The authors also proved that shocks to the beliefs had a non-negligible impact on interest and exchange

¹ Since 2004, Yahoo! Finance stock message boards have provided an option for retail investors to reveal their sentiment from among five categories: "Strong Buy", "Buy", "Hold", "Sell", and "Strong Sell".

rates. Their study showed explicitly that market sentiment reflected in individual text messages matters for asset prices.

Siganos et al. [Siganos et. al 2014] used as an alternative measure of sentiment, Facebook's Gross National Happiness Index (FGNHI) developed by Facebook's data team based on positive and negative words in status updates. They found a positive relation between sentiment on Facebook and stock market returns and a negative relation to trading volume and volatility.

The impact of social media on the capital market literature overview can be found in [Bukovina 2016].

In 2006, in addition to social media, forums and newsletters a new big data source based on the search engine queries was launched. Google Inc. released Google Trends – a public web facility that shows how often a particular search-term is entered in Google Search. As a relatively new and big data source this tool has also been recently widely used to forecast and analyse financial markets, but with mixed results.

Preis et al. [Preis et al. 2013] investigated the correlation between returns and search volume for company names but found no significant relationships. However, they provided some evidence that Google search data can be used to predict trading volume. Da et al. [Da et al. 2011] proposed a direct measure of investor attention using search data in Google and found that an increase in this measure predicted an increase in stock prices in the next two weeks and a possible price reversal within the year. Joseph et al. [Joseph et al. 2011] examined the online ticker searches in a sample of S&P 500 firms over the period 2005-2008 and found that over a weekly horizon, online search intensity reliably predicted abnormal stock returns and trading volume. Analogous conclusions came from [Bank et al. 2011], who showed evidence that an increase in search queries was linked with a growth of stock liquidity and temporarily higher returns. Preis et al. [Preis et al. 2010] built a strategy where a market portfolio was bought or sold, based on the Google search volumes. They investigated a set of 98 search terms during the period 2004 to 2011. With the keyword *debt*, this strategy outperformed the market index by 326%. In contrast, based on a panel of six years' (313 weeks) observations for 431 companies, Bijl et al. [Bijl et al. 2016] found that large numbers of searches were significant but weak predictors of future excess returns. The study [Tang, Zhu 2017] examined the day-to-day impact of an increase in abnormal Google searches on security prices within a four-week investment horizon. They analysed American depositary receipts (further: ADR) – stocks that trade in the United States but represent a specified number of shares in a foreign corporation. The authors

found only same-day increases in the prices of stocks to ADRs of both developing and developed countries. Dimpfl and Jank [Dimpfl, Jank 2016] concentrated not on prices but on volatility, and showed that a rise in the number of searches today is followed by an increase in volatility tomorrow. The authors improved volatility forecasts in-sample and out-of-sample by including search queries in autoregressive models of realized volatility.

In our study, instead of discussing the impact of GT on the stock market, we verify the hypotheses formulated in the introduction, to determine whether investor sentiment can be approximated (or replaced) with Google Trends or whether it represents a new phenomenon connected with the development of new technologies and the Internet.

3. The data

3.1. Measures of investor sentiment

The sentiment survey measures the percentage of individual investors who are classified as bullish, bearish, and neutral. Individual investors are polled from the AAII website on a weekly basis and classified based on their answer as to what direction they feel the stock market will be in next six months. The data used in this study are presented in Figure 1. Similarly to [Brown, Cliff 2004], we use the bull-bear spread (the percentage of bullish investors minus the percentage of bearish) as the sentiment measure. On average, the responses are almost evenly split with 36% of all responses bullish, 35% neutral and almost 30% bearish, and the average bull-bear spread is about 0 (see Table 1).

Table 1. Descriptive statistics of AAII sentiment measures

	Bullish	Neutral	Bearish	Bull-bear spread
Mean	0.3566	0.3485	0.2949	0.0618
Standard error	0.0051	0.0040	0.0043	0.0085
Median	0.3540	0.3388	0.2847	0.0694
Standard deviation	0.0823	0.0647	0.0690	0.1374
Sample variance	0.0068	0.0042	0.0048	0.0189
Kurtosis	-0.6509	-0.4716	0.4699	-0.2140
Skewness	0.1598	0.3112	0.7236	-0.1457
Range	0.4018	0.3126	0.3943	0.7379
Minimum	0.1775	0.2160	0.1505	-0.3517
Maximum	0.5793	0.5286	0.5448	0.3862

Source: own calculation.

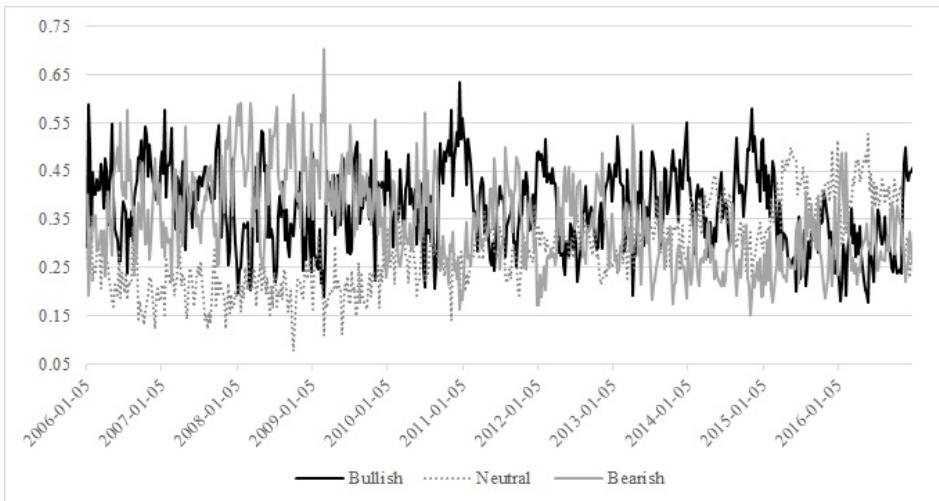


Fig. 1. The AAII survey result in the period 2012-2016

Source: own construction.

3.2. Google Trends

Google Trends is an online search tool that allows the user to see how often specific keywords, subjects and phrases have been queried over a specific period of time. Google Trends analyses the number of searches based on a query share, i.e. the total query volume for the search term within a geographic region, time and category, relative to the total number of searches within the region and time range in the category it represents. Thus, it represents the relative popularity of the keyword. The resulting numbers are then scaled within a range from 0 to 100, based on the topic's proportion to all searches in all topics.

Most of the studies mentioned in the previous paragraph measured the popularity of names or abbreviations of companies. In our study we created two sets of five keywords. The first set contains words of a negative sentiment: *crash, crisis, panic, decline, recession*, the second one those of a positive meaning: *success, confidence, rise, increase, growth*. Google Trends query was limited to the USA region and the search category was finance. The popularity of the chosen words during the study period is shown in Figure 2. We observe that the most popular among the negative words were: *crash* and *crisis*, and among the positive: *increase* and *growth*.

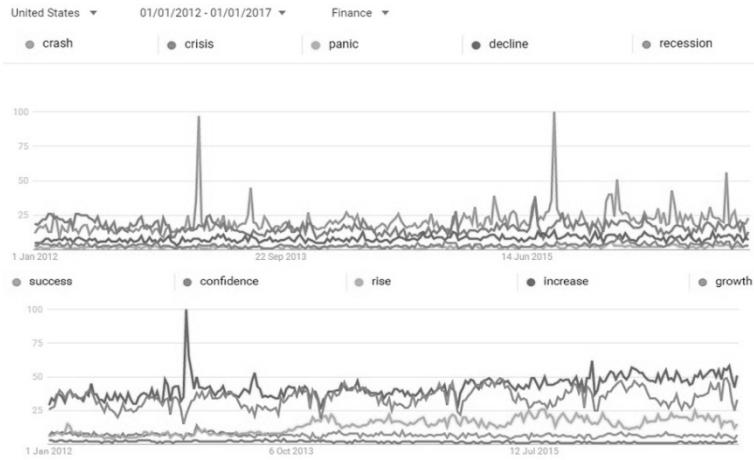


Fig. 2. The popularity of words in the period 2012-2016

Source: Google Trends.

4. The model

Let us denote by y_t a univariate time series, and by Ω_{t-1} – the set of information available up to moment $t-1$. Then the functional form of y_t can be expressed as [Laurent 2017]:

$$y_t = E(y_t | \Omega_{t-1}) + \epsilon_t,$$

where $E(\cdot | \cdot)$ denotes the conditional expectation operator, while ϵ_t is the disturbance term with $E(\epsilon_t) = 0, E(\epsilon_t \epsilon_s) = 0 \forall t \neq s$. One of the most widely useful specifications is the autoregressive: AR(k) model of the form:

$$y_t - \mu = a_0 + \sum_{i=1}^k a_i (y_{t-i} - \mu) + \sum_{i=1}^l \delta_i x_{t,i} + \epsilon_t, \quad (1)$$

where μ denotes the unconditional mean² $x_{i,t}$ denotes the value of the i -th explanatory variable at time t . In the case of most financial time-series data, ϵ_t can be decomposed in the following way:

$$\epsilon_t = z_t \sigma_t, \quad (2)$$

where $z_t \sim iid$ with mean 0 and unit variance. If σ_t^2 can be expressed as:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad (3)$$

² To estimate the model we use the G@RCH software implemented in the OxMetrics⁷ package, thus we present the specification used there to retain consistency with the results presented further.

we can say that it follows a GARCH(p, q) process (see: [Bollerslev 1986] for details). To estimate the models we used the G@RCH7.0 software provided in OxMetrics7 [Laurent 2017].

Our dependent variable (y_t) was the bull-bear spread, i.e. the difference between the percentage of bullish answers and bearish answers of the AAII survey. Hence, the higher the index, the more positive the overall investor mood. In the first step we estimated a base/reference AR-GARCH-type model (1-3) to explain linear and non-linear dependencies in investor sentiment. In order to verify hypotheses 1 and 2, we estimated the model of the same type but with additional explanatory variables. As explanatory variables we used the following numbers of search:

- all the investigated words as separate variables (separately: words of positive and negative sentiment);
- the sum of word searches from a set of selected words;
- the average search words from a set of selected words.

To verify the direction of the sentiment change together with the increase of the Google searches of the specified overtone (H1), we checked the sign and p -value of the estimated coefficients. In order to verify hypothesis H2, we assessed the estimated models based on goodness-of-fit and compared using the log-likelihood ratio test, as well as their forecasting ability. The accuracy of the forecasts was assessed based on the magnitude of the following forecasting errors:

- Mean squared error (MSE),
- Median squared error (MedSE),
- Mean error (ME),
- Mean absolute error (MAE),
- Root of the mean squared error (RMSE),
- Mean absolute percentage error (MAPE),
- Adjusted mean absolute percentage error (AMAPE),
- Percentage of Correct Signs (PCS):

$$PCS = \frac{1}{N} \max(\text{sgn}(y_{T+i} - \hat{y}_{T+i}), 0),$$

- Theil Inequality Coefficient (TIC):

$$TIC = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (y_{T+i} - \hat{y}_{T+i})^2}}{\sqrt{\frac{1}{N} \sum_{i=1}^N y_{T+i}^2} + \sqrt{\frac{1}{N} \sum_{i=1}^N \hat{y}_{T+i}^2}},$$

- Logarithmic Loss Function (LLF):

$$LLF = \frac{1}{N} \sum_{i=1}^N \ln \left(\frac{\hat{y}_{T+i}}{y_{T+i}} \right)^2.$$

In all the equations above, y_{T+i} denotes the realized value of the process, while \hat{y}_{T+i} – its forecasted value.

In order to verify hypothesis H3: *Older searches should have a weaker influence on investor sentiment than the new ones* (according to the law of small numbers), we re-estimated the models (1-3), including the lagged explanatory variables in it.

5. Results

As an opening step, we computed the descriptive statistics of the AAII time series (Table 1) and performed the test of autocorrelation and heteroscedasticity. The details are presented in Table 2.

Table 2. Results of the Box-Pierce test on raw and squared data

Statistics	Value	p-value
Box-Pierce Q-Statistics on Raw data		
Q(5)	279.7490	<0.0001
Q(10)	332.3660	<0.0001
Q(20)	363.6600	<0.0001
Q(50)	424.7790	<0.0001
Box-Pierce Q-Statistics on Squared data		
Q(5)	88.4641	<0.0001
Q(10)	125.5550	<0.0001
Q(20)	150.0060	<0.0001
Q(50)	231.8060	<0.0001

Source: own calculation.

Table 3. Base model estimates

	Coefficient	Std. error	t-value	p-value
a_0	0.0597	0.0198	3.0180	0.0028
a_1	0.6930	0.0388	17.8800	<0.0001
ω	0.0095	0.0010	9.7220	<0.0001

Source: own calculation.

Next, we estimated a base AR(1)-GARCH model describing the dynamics of the AAII time series. The AR(1) part was enough to describe linear dependencies in data. The obtained residuals did not exhibit heteroscedasticity. Log-likelihood of the model equalled 236.57. The results are presented in Table 3; then we estimated a set of models with explanatory variables, as explained in Section 2.

5.1. Interdependencies between the number of positive keywords search and investor sentiment

The results obtained for the model with all positive keywords are presented in Table 4. We put in bold the significant estimates. It turns out that the number of searches of only one of the positive keywords (“growth”) was associated with confidence growth, while the others – either with confidence decline (“success” and “increase”) or were insignificant. Next, we re-estimated the model including only significant keywords in the equation. The results are presented in Table 5. The conclusions did not change – only the number of the keyword *growth* search was associated with the increase of confidence, while the number of searches of the keywords *success* and *increase* – with its decline. A possible explanation of the phenomenon could be that the investors who lose and whose confidence as a result declines, look for information on how to become successful.

Table 4. Estimates of the model with all positive keywords

	Coefficient	Std. error	t-value	p-value
a_0	0.1672	0.0601	2.7840	0.0058
Success	-0.0070	0.0029	-2.3830	0.0179
confidence	0.0226	0.0228	0.9901	0.3231
Rise	-0.0004	0.0006	-0.5781	0.5637
Increase	-0.0004	0.0002	-2.1410	0.0332
Growth	0.0004	0.0002	2.1130	0.0356
a_1	0.6991	0.0384	18.1900	<0.0001
ω	0.0090	0.0010	9.2230	<0.0001

Source: own calculation.

Eventually we estimated the model with the mean value of all positive keyword searches (as well as with the sum of all positive keyword searches, but since according to the LL test the models have not been distinguishable, we presented only one for the sake of consistency) – see Table 6. The explanatory variable did not prove to significantly influence investor sentiment. Thus, based

on the obtained results it is hard to definitely decide about the character of the influence of the positive keyword searches on the sentiment of investors.

Table 5. Estimates of the model with three positive keywords: increase, success, growth

	Coefficient	Std. error	t-value	t-prob
a_0	0.1767	0.0597	2.9610	0.0034
Success	-0.0065	0.0030	-2.1340	0.0338
Increase	-0.0005	0.0002	-2.3680	0.0187
Growth	0.0004	0.0002	2.1410	0.0332
a_1	0.6939	0.0381	18.2100	<0.0001
ω	0.0090	0.0010	9.1560	<0.0001

Source: own calculation.

Table 6. Estimates of the model with the mean of positive word searches
as an explanatory variable

	Coefficient	Std. error	t-value	t-prob
a_0	0.1341	0.0568	2.3620	0.0189
Positive mean (M)	-0.0007	0.0005	-1.4230	0.1558
a_1	0.6945	0.0392	17.7100	<0.0001
ω	0.0095	0.0010	9.6340	<0.0001

Source: own calculation.

5.2. Interdependencies between the number of negative keyword searches and investor sentiment

In the next step we performed an analogous exercise taking into account the negative keyword searches. In Table 7 we present the estimates of the model in which all the keywords were included. Only “recession” and “panic” proved to be significant. Both coefficients were negative which is consistent with our expectations – the increase in the number of negative keyword searches is accompanied by the decline of investor sentiment.

Table 7. Model estimation with all negative word searches

	Coefficient	Std. error	t- statistics	p-value
a_0	0.1905	0.0426	4.4760	<0.0001
Crisis	-0.0003	0.0004	-0.8132	0.4168
Panic	0.0029	0.0063	0.4561	0.6487
Decline	0.0000	0.0010	-0.0053	0.9958
Recession	-0.0104	0.0031	-3.4020	0.0008
Crash	-0.0007	0.0002	-3.4670	0.0006
a_1	0.6883	0.0442	15.5800	<0.0001
ω	0.0087	0.0009	10.0400	<0.0001

Source: own calculation.

To check the validity of the conclusions we re-estimated the model considering only two keywords that proved to be significant in the previous case, see Table 8. The conclusions did not change: the increase in the number of the two keywords searches is accompanied with the decline of investor sentiment.

Table 8. Model estimation with significant negative word searches

	Coefficient	Std. error	t-value	p-value
a_0	0.1717	0.0292	5.8740	0.0000
Recession (M)	-0.0113	0.0030	-3.7030	0.0003
Crash (M)	-0.0006	0.0002	-3.3280	0.0010
a_1	0.6952	0.0405	17.1800	<0.0001
ω	0.0088	0.0009	9.9190	<0.0001

Source: own calculation.

Finally, we estimated two additional models: with the sum and mean of all negative keyword searches as explanatory variables (see Tables 9 and 10). In both cases the explanatory variable was significant and the coefficient was negative, which suggests that the increase of the negative keyword searches is accompanied by the decline of sentiment.

Table 9. Model estimation with the sum of negative word searches

	Coefficient	Std. error	t-value	p-value
a_0	0.1937	0.0393	4.9350	0.0000
Sum negative (M)	-0.0033	0.0009	-3.6470	0.0003
a_1	0.6795	0.0403	16.8700	0.0000
ω	0.0090	0.0009	9.8840	<0.0001

Source: own calculation.

Table 10. Model estimation with the mean of negative word searches

	Coefficient	Std. error	t-value	p-value
a_0	0.1937	0.0393	4.9350	0.0000
Mean negative (M)	-0.0007	0.0002	-3.6470	0.0003
a_1	0.6795	0.0403	16.8700	0.0000
ω	0.0090	0.0009	9.8850	0.0000

Source: own calculation.

5.3. Influence of past searches on investor sentiment – testing the law of small numbers

To verify the hypothesis about the weaker influence of past searches on current investor sentiment (according to the law of small numbers), we estimated the model with lagged keywords as explanatory variables. In Table 11 we present the results obtained for models with positive and negative keywords. We found no significant influence of past searches on current sentiment, although we need to stress the fact that the sentiment is quoted on a weekly basis, while the searches are performed daily. In the previous case we selected the searches performed within the week, while in this case we used the searches performed a week ago before the sentiment was announced. This means that the “oldest” search can be lagged two weeks from the sentiment announcement.

Table 11. Estimates of the models with lagged positive and negative keyword searches

	Coefficient	Std. error	t-value	p-value
Model with lagged positive keyword searches				
a_0	0.0889	0.0569	1.5630	0.1192
lagGrowth (M)	0.0000	0.0002	-0.1278	0.8984
lagSuccess (M)	0.0035	0.0024	1.4710	0.1425
lagConfidence (M)	-0.0535	0.0281	-1.9010	0.0584
lagRise (M)	-0.0004	0.0005	-0.8346	0.4047
lagIncrease (M)	0.0000	0.0002	0.0023	0.9982
a_1	0.7077	0.0432	16.4000	0.0000
ω	0.0094	0.0009	10.2600	0.0000
Model with lagged positive keyword searches				
a_0	0.0564	0.0421	1.3410	0.1811
lagCrash (M)	0.0000	0.0002	0.0181	0.9856
lagCrisis (M)	-0.0002	0.0004	-0.5031	0.6154
lagPanic (M)	0.0003	0.0062	0.0551	0.9561
lagDeclin (M)	-0.0002	0.0012	-0.1863	0.8524
lagRecession (M)	0.0048	0.0033	1.4430	0.1503
a_1	0.6941	0.0447	15.5200	0.0000
ω	0.0095	0.0010	9.5370	0.0000

Note: Although no explanatory variable proved to be significant, both models explained all linear and non-linear dependencies in the data.

Source: own calculation.

5.4. Comparison of the models – the log-likelihood ratio test

In the last step we compared the models using the log-likelihood ratio test. In each case we tested the hypothesis whether the extended model was significantly better than the base one. The results are presented in Table 12. Based on the value of the LL ratio we can say that the model including all negative keywords, as well as only “recession” and “crash” outperformed all the remaining ones. The models with positive keywords also performed better than the base one (apart from the model with the mean value of positive keyword searches). However, due to the fact that the sign of the explanatory variables differed depending on the searched keyword, we assess the models as useless for the purpose of the possible prediction of investor sentiment changes.

Table 12. Summary of models

Model	LL-ratio	p-value
Simple (base) model	--	--
Positive – all keywords	16.61	0.005302
Positive – growth, success, increase	14.972	0.001841
Positive – mean	1.928	0.164977
Positive – sum	1.928	0.164977
Negative – all keywords	21.478	0.000658
Negative – recession and crash	20.322	0.000040
Negative – mean	12.738	0.000358
Negative – sum	12.738	0.000358

Source: own calculation.

5.5. Comparison of the models based on their forecasting ability

In addition to the log-likelihood ratio test, we also compared the models taking into account their ability to forecast sentiment. We compared the base model with the models including keywords of negative sentiment. We present the results in Table 13 (positive keywords) and Table 14 (negative keywords).

In the case of the model with positive keywords and accuracy of forecasting the conditional mean, the model with “growth, confidence and success” minimizes three errors (Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE)) and the Theil coefficient, as well as maximizes the percentage correct signs. When it comes to conditional variance, which is assumed constant, the base model and the models with (insignificant) sum and mean minimize three criteria.

Table 13. Comparison of the models with positive keywords
based on their predictive availability – prediction errors

	Base model		All positive key-words		Growth, suc-cess, increase		Mean of posi-tive searches		Sum of positive searches	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
MSE	0.0056	0.0002	0.0052	0.0002	0.0051	0.0002	0.0055	0.0002	0.0055	0.0002
MedSE	0.0034	0.0001	0.0039	0.0001	0.0034	0.0001	0.0029	0.0001	0.0029	0.0001
ME	-0.0100	0.0036	0.0096	0.0031	0.0060	0.0030	-0.0071	0.0026	-0.0071	0.0026
MAE	0.0625	0.0108	0.0599	0.0103	0.0593	0.0103	0.0614	0.0103	0.0614	0.0103
RMSE	0.0750	0.0144	0.0721	0.0145	0.0715	0.0145	0.0743	0.0144	0.0743	0.0144
MAPE	.NaN	78.8500	.NaN	23.2100	.NaN	23.4000	.NaN	24.5200	.NaN	24.5200
AMAPE	.NaN	0.5015	.NaN	0.5181	.NaN	0.5169	.NaN	0.5106	.NaN	0.5106
PCS	0.7000	.NaN	0.7667	.NaN	0.7333	.NaN	0.6667	.NaN	0.6667	.NaN
TIC	0.3692	0.4868	0.3683	0.5094	0.3650	0.5074	0.3830	0.4963	0.3830	0.4963
LL	.NaN	5.8710	.NaN	5.4560	.NaN	5.4000	.NaN	5.5620	.NaN	5.5620

Note: the minimum values of each error are in bold.

Source: own calculation.

Table 14. Comparison of the models with negative keywords
based on their predictive availability – prediction error

	Base model		All negative keywords		Crash + recession		Mean of negative searches		Sum of negative searches	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
MSE	0.0056	0.0002	0.0061	0.0002	0.0061	0.0002	0.0054	0.0002	0.0054	0.0002
MedSE	0.0034	0.0001	0.0041	0.0001	0.0041	0.0001	0.0031	0.0001	0.0031	0.0001
ME	-0.0100	0.0036	-0.0016	0.0035	0.0000	0.0035	-0.0083	0.0031	-0.0083	0.0031
MAE	0.0625	0.0108	0.0640	0.0104	0.0643	0.0104	0.0608	0.0103	0.0608	0.0103
RMSE	0.0750	0.0144	0.0780	0.0146	0.0781	0.0146	0.0732	0.0145	0.0732	0.0145
MAPE	.NaN	78.8500	.NaN	22.2900	.NaN	22.3600	.NaN	23.2900	.NaN	23.2900
AMAPE	.NaN	0.5015	.NaN	0.5237	.NaN	0.5233	.NaN	0.5176	.NaN	0.5176
PCS	0.7000	.NaN	0.7000	.NaN	0.7000	.NaN	0.7000	.NaN	0.7000	.NaN
TIC	0.3692	0.4868	0.3947	0.5193	0.3918	0.5185	0.3788	0.5085	0.3788	0.5085
LL	.NaN	5.8710	.NaN	5.3820	.NaN	5.3880	.NaN	5.4630	.NaN	5.4630

Note: the minimum values of each error are in bold.

Source: own calculation.

In the case of the model with negative keywords – if we wanted to minimize the mean error of sentiment forecast, we should apply the model with the sum or mean value of negative searches: Mean Squared Error (MSE), Median Squared Error (MedSE), Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are minimized. When it comes to forecasting of variance (which is assumed constant in the model), the results are not so obvious. The model with the sum or mean of all negative keyword searches minimizes three errors: the median squared error (MedSE), the mean squared error (MSE) and the mean absolute error (MAE). The root mean squared error, adjusted mean absolute percentage error, as well as the Theil coefficient (TIC) are the lowest in the case of the base model.

6. Conclusion and future research

The aim of the article was to verify whether the number of selected keyword searches can reflect the dynamics of individual investor sentiment. Hence we collected five keywords of positive sentiment, as well as five keywords of negative sentiment and checked the number of their searches using the Google Trend tool. As a measure of individual investor sentiment we used the bull-bear spread of the AAII index – a survey-based index of American individual investor sentiment. We estimated the model of the individual investor sentiment dynamics using the ARMA-GARCH type models. As the data is published weekly, the dynamics of the sentiment was not very volatile and the AR(1) model was enough to explain all linear and non-linear dependencies in the data. In the next step we re-estimated the model, including the set of positive and negative keywords as explanatory variables, separately. The results of our study support the following hypotheses, describing the sentiment of investors from the behavioural finance point of view:

- Negative keyword searches reflect stronger investor sentiment dynamics than those positive: the sign of the coefficient is always negative, suggesting that the decrease of investor sentiment and increase of negative keyword searches appear at the same time. This result partially confirms the risk aversion of investors and their tendency to react stronger to negative news than to the positive ones (e.g. the leverage effect).
- The law of small numbers is visible in investor sentiment dynamics – the impact of past keyword searches is insignificant, while the impact of current (within a week) searches is significant.

We compared the models in-sample using the log-likelihood ratio test. The results strongly support the superiority of the model with negative keywords. The out-of-sample comparison of the ability of the models to forecast investor sentiment also suggests the superiority of the models including keyword searches as explanatory variables over the base AR(1) model. However, the model with negative keywords only does not seem to outperform the model with positive keywords, although it was obviously better when it comes to in-sample comparison. This suggests that when one wants to model the sentiment of investors with the help of keyword searches, we should include both types of keywords – the positive and the negative ones in the model. The choice of the keywords is still an open question.

Acknowledgements

The authors would like to thank the participants of the conference **ICOFEP**: International Conference on Finance and Economic Policy (Poznań, Poland, 27-28.04.2017) for a fruitful discussion and all the valuable comments on the first version of this paper, and also the participants of the seminar of the Applied Mathematics Department (especially Professor Małgorzata Doman) for all the remarks considering the model section.

Bibliography

- Antweiler W., Frank, M.Z. (2004). *Is all that talk just noise? The information content of internet stock message boards.* The Journal of Finance. Vol. 59. No. 3, pp. 1259-1294.
- Bank M., Larch M., Peter G. (2011). *Google search volume and its influence on liquidity and returns of German stocks,* Financial Markets and Portfolio Management. Vol. 25. No. 3, pp. 239-264.
- Baker M., Wurgler J. (2006). *Investor sentiment and the cross-section of stock returns.* The Journal of Finance. Vol. 61. No. 4, pp. 1645-1680.
- Bijl L., Kringshaug G., Molnár P., Sandvik E. (2016). *Google searches and stock returns.* International Review of Financial Analysis. Vol. 45, pp. 150-156.
- Bollen J., Mao H., Zeng X. (2011). *Twitter mood predicts the stock market.* Journal of Computational Science. Vol. 2. No. 1, pp. 1-8.
- Bollerslev T. (1986). *Generalized autoregressive conditional heteroscedasticity.* Journal of Econometrics. Vol. 31. No. 3, pp. 307-327.
- Brown G.W., Cliff M.T. (2004). *Investor sentiment and the near-term stock market.* Journal of Empirical Finance. Vol. 11. No. 1, pp. 1-27.
- Brown G.W., Cliff M.T. (2005). *Investor sentiment and asset valuation.* The Journal of Business. Vol. 78. No. 2, pp. 405-440.

- Bukovina J. (2016). *Social Media and Capital Markets. An Overview*. Procedia-Social and Behavioral Sciences. Vol. 220, pp. 70-78.
- Campbell J.Y., Hentschel L. (1992). *No news is good news: An asymmetric model of changing volatility in stock returns*. Journal of Financial Economics. Vol. 31. No. 3, pp. 281-318.
- Da Z., Engelberg J., Gao P. (2011). *In search of attention*. The Journal of Finance. Vol. 66. No. 5, pp. 1461-1499.
- De Long J.B., Shleifer A., Summers L.H., Waldmann R.J. (1990). *Noise trader risk in financial markets*. Journal of Political Economy. Vol. 98. No. 4, pp. 703-738.
- Dimpfl T., Jank S. (2016). *Can internet search queries help to predict stock market volatility?* European Financial Management. Vol. 22. No. 2, pp. 171-192.
- Fisher K.L., Statman M. (2000). *Investor sentiment and stock returns*. Financial Analysts Journal. Vol. 56. No. 2, pp.16-23.
- Hibbert A.M., Daigler R.T., Dupoyet B. (2008). *A behavioral explanation for the negative asymmetric return–volatility relation*. Journal of Banking & Finance. Vol. 32. No. 10, pp. 2254-2266.
- Joseph K., Wintoki M.B., Zhang Z. (2011). *Forecasting abnormal stock returns and trading volume using investor sentiment: evidence from online search*. International Journal of Forecasting. Vol. 27. No. 4, pp. 1116-1127
- Kim S.H., Kim D. (2014). *Investor sentiment from internet message postings and the predictability of stock returns*. Journal of Economic Behavior & Organization. Vol. 107, pp. 708-729.
- Laurent S. (2017). G@RCH 7.0 Help, file:///C:/program%20files/oxmetrics7/doc/g@rch/index.html.
- Meinusch A., Tillmann P. (2017). *Quantitative Easing and Tapering Uncertainty: Evidence from Twitter*. Joint Discussion Paper Series in Economics. No. 09-2015.
- Oliveira N., Cortez P., Areal N. (2017). *The impact of microblogging data for stock market prediction: using Twitter to predict returns, volatility, trading volume and survey sentiment indices*. Expert Systems with Applications. Vol. 73, pp. 125-144.
- Piñeiro-Chousa J.R., López-Cabarcos M.Á., Pérez-Pico A.M. (2016). *Examining the influence of stock market variables on microblogging sentiment*. Journal of Business Research. Vol. 69. No. 6, pp. 2087-2092.
- Preis T., Reith D., Stanley H.E. (2010). *Complex dynamics of our economic life on different scales: Insights from search engine query data*. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. Vol. 368. No. 1933, pp. 5707-5719.
- Preis T., Moat H.S., Stanley H.E. (2013). *Quantifying trading behavior in financial markets using Google Trends*. Scientific reports 3, srep01684.
- Schmeling M. (2009). *Investor sentiment and stock returns: Some international evidence*. Journal of Empirical Finance. Vol. 16. No. 3, pp. 394-408.
- Siganos A., Vagenas-Nanos E., Verwijmeren P. (2014). *Facebook's daily sentiment and international stock markets*. Journal of Economic Behavior & Organization. Vol. 107, pp. 730-743.

-
- Sun A., Lachanski M., Fabozzi F.J. (2016). *Trade the tweet: Social media text mining and sparse matrix factorization for stock market prediction*. International Review of Financial Analysis. Vol. 48, pp. 272-281.
- Tang W., Zhu L. (2017). *How security prices respond to a surge in investor attention: Evidence from Google Search of ADRs*. Global Finance Journal. Vol. 33, pp. 38-50.
- Tversky A., Kahneman D. (1975). *Judgment under Uncertainty: Heuristics and biases*. In: *Utility, Probability, and Human Decision Making*. Springer Netherlands, pp. 141-162.