

# Can Retail Investor Attention enhance Market Efficiency? Insights from Search Engine Data

**Konstantin Storms\*** **Julia Kapraun\*** **Markus Rudolf\***

November 3, 2015

## **Abstract**

This study provides novel insights to the ongoing debate how market efficiency is challenged by investor behavior. Applying search engine data we find that retail investor attention can enhance market efficiency. High attention is associated with better incorporation of idiosyncratic stock information, which we interpret as improved pricing efficiency. This effect is even more pronounced in bullish markets. In bearish markets, however, retail investor attention leads to a deterioration of pricing efficiency, which might be explained with herding behavior. Our evidence holds for a broad sample of European and US stocks.

*JEL classification:* G02, G12, G14

*Keywords:* market efficiency, investor attention, search engine volume

\*All authors are from WHU – Otto Beisheim School of Management, Campus Vallendar, Burgplatz 2, 56179 Vallendar, Germany. Corresponding author is Konstantin Storms, please contact at +49 261 6509-398, Konstantin.Storms@whu.edu, [www.whu.edu/finance](http://www.whu.edu/finance). The authors would like to thank Walter Herzog, Maximilian Müller, Marcel Omachel, Mei Wang and Kai Winselmann for helpful comments and constructive feedback.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Related literature</b>	<b>4</b>
<b>3</b>	<b>Methodology and hypotheses</b>	<b>6</b>
3.1	Google Search volume as retail investor attention proxy . . . . .	6
3.2	Measuring investor sentiment . . . . .	8
3.3	Investor attention as driver of trading activity . . . . .	9
3.4	Investor attention as driver of market efficiency . . . . .	11
<b>4</b>	<b>Data</b>	<b>13</b>
<b>5</b>	<b>Results</b>	<b>15</b>
5.1	Trading activity . . . . .	15
5.1.1	Trading volume . . . . .	15
5.1.2	Return . . . . .	20
5.2	Market efficiency . . . . .	21
5.2.1	Incorporation of idiosyncratic information . . . . .	22
5.2.2	Return predictability . . . . .	24
<b>6</b>	<b>Conclusion</b>	<b>26</b>
	<b>References</b>	<b>27</b>
	<b>Appendix</b>	<b>30</b>

# 1 Introduction

The efficient market hypothesis (EMH) asserts that all newly available information is immediately reflected in asset prices, implying that capital markets are informationally efficient (Fama, 1970). Market efficiency requires rational investors: they discover a piece of information, derive its meaning for the fundamental value of the corresponding security, act accordingly and thereby bring prices to a level reflecting available information (Merton, 1987). However, several studies oppose that investors are limited in their cognitive capabilities and focus their scarce resource attention. For instance, when the New York Times reported the discovery of a new cancer-curing drug by EntreMed on the front page of its Sunday issue, the EntreMed stock generated a return of more than 300% the next Monday and long-term remained on a higher price level. But all this was stale news and had been published in Nature and other newspapers several months before (Huberman and Regev, 2001). Investors simply did not pay attention. This impressive incident raises the question to what extent markets can be efficient if investors are inattentive, or more precisely, if investor attention enhances market efficiency.

In order to capture investor attention, several measures have been proposed so far. Most recently, Da et al. (2011) introduce a direct attention measure and emphasize that searching something on the internet is a definite sign of paying attention. They explore that Google Search volume (GSV) for US stocks' ticker correlates with other investor attention proxies and predicts short-term price movements. As institutional investors rather rely on more sophisticated tools to search for stock related information, GSV mainly captures the attention of retail investors. Retail investors are typically described as uninformed noise traders (e.g. Odean, 1998; Grinblatt and Keloharju, 2000) and thus could cause prices to deviate from fundamentals by trading on non-information. This would imply that their attention corrupts market efficiency.

In our study we take a more differentiated view on this relation and question, if there are certain circumstances under which retail investor attention can actually enhance market efficiency. Our contribution to related research is twofold. First, we build on Da et al. (2011) and develop a novel Google Search term approach using firm name and the word "stock" in local language in order to capture retail investor attention. This method explains more variance in trading volume than alternative search terms including firm name and ticker. Second, we are, to our knowledge, the first ones to highlight the relation between retail investor attention and an individual security's pricing efficiency in different market states. We reveal that high retail investor attention leads to better

incorporation of an individual stock’s idiosyncratic information. Prices are then to a larger extent based on firm-specific information rather than being driven by market-wide factors, which we interpret as improved pricing efficiency. However, retail investors may base their trading decisions on beliefs rather than facts (often referred to as investor sentiment), which can lead to mispricing (De Long et al., 1990; Da et al., 2015). This effect is particularly strong in times of financial uncertainty or if markets decrease. To account for this we distinguish different market states and restate that in times of bullish markets and low financial uncertainty, the enhancement of efficiency through attention is even more pronounced. But in times of bearish markets, attention is associated with a deterioration of pricing efficiency. This might be explained with well-known behavioral biases of retail investors. Rather than acting according to true changes in fundamentals, retail investors could behave according to behavioral heuristics like herding, leading to pricing inefficiency. This behavior leads to stronger co-movement of returns and, ultimately, impairs pricing efficiency. To stress robustness of our insights, we additionally illustrate that in bullish markets retail investor attention mitigates a market anomaly: it reduces return predictability.

Our results, stemming from a differentiated view on the relation between retail investor attention and market efficiency, provide novel insights to the ongoing debate how market efficiency is challenged by behavioral patterns. The remainder of this study is structured as follows: first we present related investor attention literature. Then we develop hypotheses and explain our methodology. In the next sections we introduce the data set, provide main empirical results and conclude with a summary of our findings.

## 2 Related literature

The literature documents broad evidence for limited investor attention. When forming portfolio allocation decisions, investors are typically not aware of all securities and consider a subset of assets (Merton, 1987). Investors only notice “stocks on their radar screens” (Daniel et al., 2002, p. 144) and therefore are buyers of “attention grabbing stocks” (Barber and Odean, 2008).

Investor attention can be captured in manifold ways. Fang and Peress (2009) measure investor attention by media coverage and advocate, the more stocks are present in the news, the higher should be the attention. Stocks with no media coverage and, hence, with low attention earn higher abnormal returns. If the absence of abnormal returns is considered efficient, this study provides first evidence that attention enhances market

efficiency. Dellavigna and Pollet (2009) postulate that attention is lower on Fridays. Earnings announcements on Fridays experience lower immediate price responses on the announcement day and lead to a higher drift in the direction of the earnings surprise, which is a violation of market efficiency. A similar effect can be observed, when several announcements compete for attention (Hirshleifer et al., 2009).

Drake et al. (2012) apply a direct attention measure and use Google Search Volume (GSV) to explain returns on earnings announcement days. They illustrate that if search volume prior announcement is high, the abnormal return on the announcement day is lower. They interpret this weaker stock price reaction as information regarding an earnings surprise being partially preempted through attentive investors. Fricke et al. (2014) observe that GSV not only reduces abnormal returns on the earnings announcement day, but also mitigates the post-earnings announcement drift. These two studies provide preliminary evidence that investor attention, as measured by GSV, increases market efficiency by extenuating a market anomaly. Vozlyublennaia (2014) conducts a study for market indexes because she assumes that retail investors typically invest into broad index funds rather than buying individual securities. According to her study GSV influences performance of US stocks and commodity indexes and drives returns in the short-run. Additionally, higher GSV reduces return predictability of the index, which Vozlyublennaia emphasizes as an enhancement of market efficiency. Her study, however, does not provide any answer, whether retail investor attention towards individual stocks can enhance the pricing efficiency of that security.

Very few studies conduct search volume investigations in a European context. Bank et al. (2011) analyze a sample of German stocks and find that GSV improves liquidity and increases short-term buying pressure. Tickers are not commonly used in Germany in order to identify stocks. For this reason, the authors decide to use firm name as search term, regardless of some shortcomings of this method. Also Aouadi et al. (2013) use GSV for firm names and show that it is associated with increased liquidity and volatility in a sample of French CAC 40 stocks. Siganos (2013) examines how GSV can predict prior M&A announcement returns for UK firms, which he interprets as investors expecting the merger and trading on that insight.

Recent research does not explicitly focus on an individual security's pricing efficiency resulting from retail investor attention. Furthermore, especially for a European application, using firm name as search term seems not ideal, as we will explain later. We will fill this gap and present an alternative search term approach as well as explicitly focus on pricing efficiency of individual securities in different market states.

### 3 Methodology and hypotheses

#### 3.1 Google Search volume as retail investor attention proxy

Recent research applying Google Search volume in a European context propose to use firm name as the adequate search term (e.g. Bank et al., 2011; Jacobs and Weber, 2011; Aouadi et al., 2013). However, searching firm name on Google can be motivated by various reasons. For instance, when searching “Deutsche Bank” on the internet, users are most likely looking for opening hours of their local branch or the log-in page to the online banking system. Only few search hits for firm name might be motivated by looking for financial information and stock related data. Also the “related search terms” option in Google reveals that most users are interested in the online banking tool (“meine Deutsche Bank”), branch opening hours (“Öffnungszeiten Deutsche Bank”) or career opportunities (“Deutsche Bank Karriere”). Search volume for raw firm name thus seems to be quite arbitrary, if it is ought to capture pure investor attention. For this reason we propose a new search term methodology and use firm name in combination with the word “stock” in local language. We use local language as we assume that mainly domestic retail investors search for domestic stocks with very few exceptions.

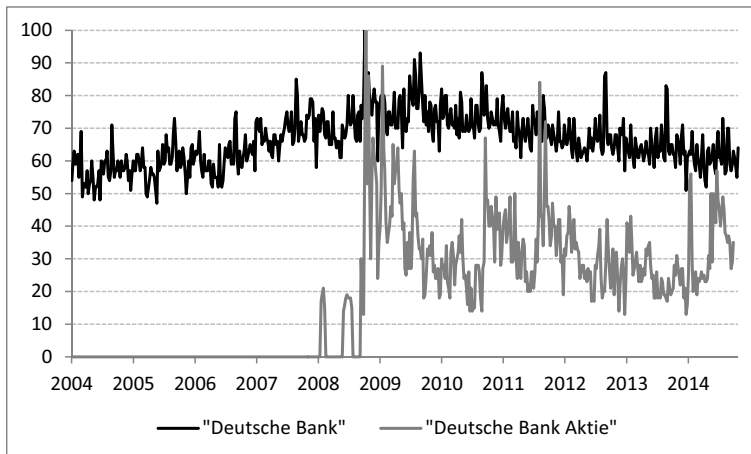


Figure 1: Google Search Volume “Deutsche Bank”

Let us illustrate our approach for the stock of Deutsche Bank. As shown in Figure 1, search volume for “Deutsche Bank” is constantly high over time, with relatively low variation, while search volume for “Deutsche Bank Aktie” (“Aktie” is “stock” in German language) is zero before 2008 and fluctuates heavily over time. This impression is confirmed when consulting the statistics in Table 1. Google Trends provides a search

volume score, which is scaled from 0 to 100. Thereby 100 represents the maximum search volume over the corresponding period. We observe a mean Google Search volume score of 66 with low standard deviation of 8 for “Deutsche Bank” and a mean score of 19 with comparatively high standard deviation of 19 for “Deutsche Bank Aktie”. Correlation between these two search terms is positive, but not very high ( $\rho = 0.49$ ).

Table 1: Google Search volume “Deutsche Bank” statistics

Search term	Statistics			Quartiles					Pearson Correlation	
	N	Mean	St.Dev.	Min	0.25	Med	0.75	Max	(1)	(2)
(1) “Deutsche Bank”	561	66.40	8.05	47	61	66	71	100	1.00	
(2) “Deutsche Bank Aktie”	561	19.02	19.02	0	0	20	32	100	0.49	1.00

Notes: Sample consists of 561 weeks with Google Search volume score available

Several related search terms are suggested by Google when typing in firm name and “stock” in local language. We rely on the search term that produces the highest relative search volume. This approach does not bias our results, because the suggested terms are highly correlated by definition.<sup>1</sup> We come up with a final selection of search terms for “stock” as presented in Table 2.

Table 2: Search term for “stock” in local language

Country	Search term
Belgium	Action
France	Action
Germany	Aktie
Italy	Azioni
Netherlands	Koers
Spain	Acciones
Switzerland	Aktie
United Kingdom	Stock
United States	Stock

We manually downloaded weekly Google Search data from the Google Trends website<sup>2</sup> over the course of October and November 2014, using search terms as described, e.g. “Deutsche Bank Aktie”, “Apple Stock”, “Fiat Azioni” or “Sanofi Action”. For those firms

<sup>1</sup>Two exceptions are Belgium and Switzerland as multilingual countries. We use the French word “Action” and the German word “Aktie” for Belgium and Switzerland respectively, as it produces the highest search volume for most stocks. This, however, might bias our results, as it excludes Flemish, or French and Italian speaking domestic investors.

<sup>2</sup>[www.google.com/trends](http://www.google.com/trends)

with weekly data available, we additionally download daily data by limiting the time period to three months per request. Google scales the daily search volume data relative to the maximum volume within this three month period. In order to compare GSV scores across quarters, we rescale the daily GSV score using the weekly score with:

$$GSV_{Scaled}^{Daily} = \frac{GSV_{Unscaled}^{Daily} \times GSV^{Weekly}}{100} \quad (1)$$

Following Da et al. (2011), we calculate abnormal search volume (ASV) with:

$$ASV_t = \ln[GSV_t] - \ln[Med(GSV_{t-1}, \dots, GSV_{t-k})] \quad (2)$$

where k is 8 weeks for weekly GSV scores or 56 days (i.e., 8 weeks times 7 week days) for daily GSV scores.

### 3.2 Measuring investor sentiment

Google search volume captures the attention of retail investors (Da et al., 2011). Retail investors are often considered uninformed noise traders (e.g. Dimpfl and Jank, 2015; Grinblatt and Keloharju, 2000; Odean, 1998) who base investment decisions on sentiment, i.e., believes rather than facts (De Long et al., 1990; Da et al., 2015). Therefore, we will control for sentiment in our analyses. The level of investor sentiment, or the degree to which decisions are based on believes, can be measured by various proxies. Those include e.g. trading volume, initial public offering (IPO) first day returns or mutual fund flows (see Baker and Wurgler, 2007, for a comprehensive review). A common measure of investor sentiment are market volatility indexes, with the US CBOE Volatility Index (VIX) as most prominent example. The index’s level, denoted in percent, expresses the consensus view of the market about future expected volatilities. As increased volatility means increased risk, the VIX is often called the “investor fear gauge”. The greater the VIX, the greater the fear of the market (Whaley, 2000).

In our study we use the most applicable market volatility index for each country as presented in Table 10 (in the appendix) to construct an investor sentiment measure. Market volatility indexes are highly volatile as “normal” or average levels change quite frequently (see Figure 2 for the VIX). Between January 2004 and February 2015, the US VIX ranged between 9.9% and 80.9%, while 50% of the closing index level lies in a range between 13.4% and 22.4% (see Table 10 in the appendix).

To account for these heavy fluctuations in the index, we measure sentiment by a  $Fear_{it}$



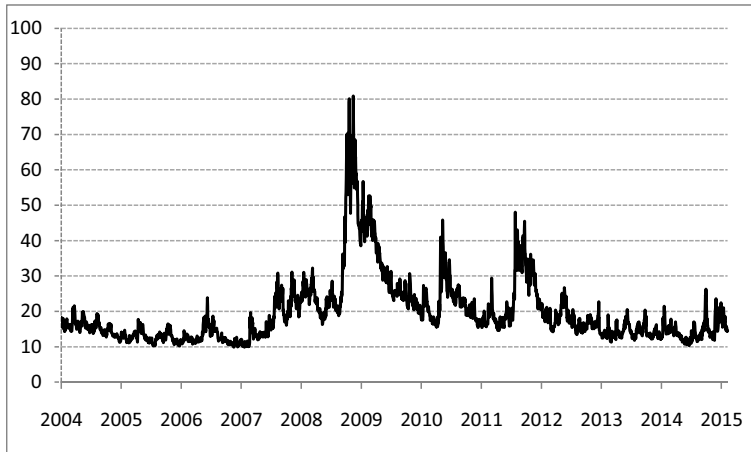


Figure 2: VIX index

dummy variable, which takes the value of 1, if the corresponding market volatility index (VI) increases and zero otherwise:

$$Fear_{it} := \begin{cases} 1, & \text{if } VI_{it} > VI_{it-1} \\ 0, & \text{if } VI_{it} \leq VI_{it-1} \end{cases} \quad (3)$$

### 3.3 Investor attention as driver of trading activity

Similar to Bank et al. (2011) and also Aouadi et al. (2013), who investigate the relation of GSV and stock liquidity, we first test, whether our modified attention proxy can explain trading activity. If investors are attentive towards a certain stock and make investment decisions, attention should lead to trading volume. Therefore we hypothesize:

**Hypothesis 1a:** *High retail investor attention as captured by Google Search volume for firm name and “stock” in local language explains increasing trading volume.*

To confirm hypothesis 1a we use ordinary least squares (OLS) regression with firm and year fixed effects as well as robust standard errors. We regress:

$$AVO_{it} = \alpha_i + \beta_{1i}AVO_{it-1} + \beta_{2i}ASV_{it} + \sum_{k=1}^K \beta_{3ki}Control_{kit} + \epsilon_{it} \quad (4)$$

$AVO_{it}$  is abnormal trading volume calculated with current trading volume (VO) relative

to median trading volume of the last 40 trading days (i.e., 8 weeks) with:

$$AVO_{it} = \ln [VO_{it}] - \ln [Med(VO_{it-1}, \dots, VO_{it-40})] \quad (5)$$

Weekly AVO is average daily abnormal trading volume in the corresponding week.  $AVO_{it-1}$  is abnormal trading volume lagged by one week,  $ASV_{it}$  is abnormal Google search volume for firm  $i$  in week  $t$ . Our  $K$  control variables include  $Vol_{it}$ , i.e., standard deviation of last 180 trading days' returns,  $Size_{it}$  i.e., logarithm of total market capitalization,  $BTM_{it}$ , i.e., book value of equity divided by total market capitalization,  $InstOwn_{it}$ , i.e., fraction of shares held by institutional owners (investment companies and pension funds) and  $BASpread_{it}$ , i.e., bid ask spread, calculated as difference between ask (PA) and bid price (PB), deflated by the mid-point of ask and bid price. If attention captured by Google search volume drives abnormal trading volume (Hypothesis 1a), we should observe  $\beta_2 > 0$ .

After having analyzed whether attention leads to trades, we next investigate, if attention allows us to predict returns as well. It is per se not clear, if attention leads to a sale or a buy of an asset. However, as explained earlier, retail investors are sensitive towards sentiment. If the market is in fear, retail investors are likely intimidated and tend to sell securities. Therefore, if fear increases, we expect investors to panic and to sell. Therefore, we hypothesize:

**Hypothesis 1b:** *If the market is in "fear" the stock return decreases. This effect is more pronounced if retail investor attention is high.*

To validate hypothesis 1b we regress:

$$r_{it} = \alpha_i + \beta_{1i}Fear_{it} + \beta_{2i}ASV_{it} + \beta_{3i}ASV_{it} \times Fear_{it} + \sum_{k=1}^K \beta_{4ki}Control_{kit} + \epsilon_{it} \quad (6)$$

If returns decrease whenever the market is in fear (Hypothesis 1b), we should find  $\beta_1 < 0$ . We expect this effect to be even more pronounced, if retail investor attention is high and hence should observe  $\beta_3 < 0$ .

### 3.4 Investor attention as driver of market efficiency

After confirming that investor attention is associated with trading activity, we investigate if attention impacts market efficiency as well. When defining market efficiency, we consider efficiency of prices of an individual security, which is often referred to as “pricing efficiency” (e.g. Chang et al., 2014). It is hard to believe, that markets are always and at any time fully efficient. Market efficiency is not necessarily a binary state or a yes/no question, with markets either being efficient or not. According to Ken French we should rather “think about it as a continuum, where a score of 100 means prices are exactly right and a score of zero means prices are exactly wrong” (Doukas et al., 2002, p. 235). But how to measure efficiency?

Chang et al. (2014) measure market efficiency gains resulting from removed short-selling constraints and apply a market-model based on Bris et al. (2007). They regress returns of an individual stock on the return of the market index and calculate the respective  $R^2$  of that regression. The  $R^2$  measures the part of variation in the dependent variable (stock return) that can be explained by the variation in the independent variable (market return). The remaining variation comes from the idiosyncratic factor or the regression residual. Therefore, a low  $R^2$  suggests that the return of the security is more driven by idiosyncratic or firm specific information rather than by market-wide information. Hence, lower  $R^2$  can be interpreted as a sign of higher pricing efficiency. We build on this insight and hypothesize that high attention is associated with higher pricing efficiency:

**Hypothesis 2a:** *High retail investor attention leads to better **incorporation of idiosyncratic stock information**, i.e., it enhances the pricing efficiency of the corresponding security.*

We verify hypothesis 2a by following Bris et al. (2007) and Chang et al. (2014) and estimate the market model using OLS for each firm  $i$  in our sample with:

$$r_{it} = \alpha_i + \beta_i mkt_{it} + \epsilon_{it} \quad (7)$$

where  $r_{it}$  is the raw return of firm  $i$  in week  $t$ ,  $mkt_{it}$  is the return of the corresponding market index in week  $t$ . For each firm we divide the sample of all firm week observations into two groups based on abnormal search volume (ASV) with:

- High attention group:  $ASV_{it} > 0$ , i.e., weeks of firm  $i$  with attention higher than

median attention of last 8 weeks

- Low attention group:  $ASV_{it} \leq 0$ , i.e., weeks of firm  $i$  with attention lower than or equal to median attention of last 8 weeks

Regressions with insignificant F-test and insignificant t-test for the  $\beta_1$  coefficient, both at 5% level, are excluded from the analysis. The delta of the  $R^2$  of the high and the low group for each firm  $i$  is calculated with:

$$delta_i = R_{High}^2 - R_{Low}^2 \quad (8)$$

We calculate the mean of all deltas  $\mu_{delta}$  across  $N$  firms and test for statistical significance by applying a t-test:

$$t = \frac{\mu_{delta}}{\sigma_{delta}/\sqrt{N}} \quad (9)$$

An overall significantly negative  $\mu_{delta}$  indicates that high investor attention is associated with less variance in stock returns that can be explained by variance in market returns, pointing at increased pricing efficiency (Hypothesis 2a).

As we assume that retail investors behave differently due to sentiment, we additionally distinguish state of the market. First, we divide the sample in two subgroups based on market return in order to assess whether the market is bullish or bearish:

- $R^2$  Bullish:  $mkt_t \geq 0$ , i.e., weeks with zero or positive market return
- $R^2$  Bearish:  $mkt_t < 0$ , i.e., weeks with negative market return

Second, we divide the sample into subgroups depending on market fear as measured by the volatility index (VI):

- $R^2$  Unfearful:  $Fear_t = 0$ , i.e., weeks with constant or decreasing level of VI
- $R^2$  Fearful:  $Fear_t = 1$ , i.e., weeks with increasing level of VI

As a robustness check we apply an additional efficiency measure. If markets are efficient, we should not observe any market anomalies. Therefore, another typical way of testing whether a certain event or treatment (e.g high attention on a stock) enhances efficiency is to assess, whether it extenuates a market anomaly.

A prominent market anomaly is return predictability, which is a severe violation of the weak form of market efficiency. If markets are efficient and all information of today is incorporated in today’s security prices, price changes or returns should be unpredictable based on historic returns (Malkiel, 2003; Fama, 1970, 1991). Hence, the reduction of return predictability would be a sign for an enhancement of market efficiency (e.g. Vozlyublenniaia, 2014). Thus we hypothesize:

**Hypothesis 2b:** *High retail investor attention reduces **return predictability**, i.e., it enhances market efficiency.*

Testing hypothesis 2b we regress:

$$r_{it} = \alpha_i + \beta_{1i}r_{it-1} + \beta_{2i}ASV_{it-1} + \beta_{3i}r_{it-1} \times ASV_{it-1} + \sum_{k=1}^K \beta_{4ki}Control_{kit} + \epsilon_{it} \quad (10)$$

Thereby,  $r_{it-1}$  is return lagged by one week. We use the same controls as for hypothesis 1a. If we observe return predictability, i.e., lagged returns allow for prediction of current returns, we should find a statistically significant  $\beta_1$ . If, however, attention in the last week was high at the same time, this effect should be less pronounced.  $\beta_3$  taking the opposite sign as  $\beta_1$  is therefore consistent with high investor attention being associated with reduced return predictability.

## 4 Data

We focus our analyses on major European economies including Belgium, France, Germany, Italy, Netherlands, Spain, Switzerland and United Kingdom as well as United States.<sup>3</sup> Thereby, we consider all stocks listed in the main stock index of the respective country as retrieved from Thomson Reuters Datastream constituent lists.

---

<sup>3</sup>Austria, Norway, Poland and Sweden omitted due to low number of available Google Search data.

## Google Search volume

Google only provides weekly search volume data if a certain threshold is reached. Therefore, we can calculate weekly ASV data for 220 stocks out of the original sample of 2,668 stocks considered, providing 58,946 firm weeks. Out of these 220 stocks, Google provides sufficient daily data for 106 firms, leading to 88,502 daily ASV scores (see Table 3).

Table 3: Google search data per country

Country	Stocks considered	Weekly ASV score available	Firm weeks with ASV score available	Daily ASV score available	Firm days with ASV score available
Belgium	127	2	334		
France	520	25	6,551	10	7,209
Germany	160	34	5,766	12	7,562
Italy	280	9	1,721	6	2,331
Netherlands	115	2	486		
Spain	107	6	1,533	3	1,967
Switzerland	211	3	655		
United Kingdom	646	17	5,613	9	8,319
United States	502	122	36,287	66	61,114
Total	2,668	220	58,946	106	88,502

Notes: Stocks considered is number of stocks for which Google data has been collected, ASV score available is number of stocks for which Google Trends provides search data, firm weeks/firm days with ASV score available is number of observations for which the corresponding Abnormal Search volume score can be calculated.

## Market data

We retrieve all market data, i.e., return index (RI), trading volume (VO), bid price (PB), ask price (PA), total market capitalization (MV), institutional ownership (NOSHIC plus NOSHPF), Book-to-market (WC03501 divided by MV) from Thomson Reuters Datastream. We use logarithmic returns with:

$$r_{it} = \ln \left( \frac{RI_{it}}{RI_{it-1}} \right) \quad (11)$$

## Summary statistics

Table 11 in the appendix provides a detailed description of all defined variables including source and calculation. The sample consisting of weekly data is described in Table 12 (in the appendix). The correlation analysis reveals that abnormal search volume for name and “stock” in local language,  $ASV_t^{Stock}$ , is indeed positively correlated with trading volume ( $\rho = 0.32$ ). The correlation is stronger than for GSV based on firm name

( $\rho = 0.14$ ) or ticker ( $\rho = 0.28$ ). The sample based on daily GSV scores, for which we gathered data for  $ASV_t^{Stock}$ , is summarized in Table 13 (in the appendix).

## 5 Results

We first verify that our investor attention proxy explains trading activity. Thereby, we show that its predictive power is higher than that of proxies based on alternative search terms. Turning to market efficiency measures, we first validate that attention increases pricing efficiency, i.e., it leads to better incorporation of firm specific information. To stress robustness we show that attention can mitigate return predictability.

### 5.1 Trading activity

#### 5.1.1 Trading volume

If investors pay attention to certain stocks, this should lead to an investment decision and ultimately to a trade. Hence, high attention is hypothesized to be associated with increased trading volume. Results of our trading volume analysis are presented in Table 4.

As we observe in column (1), average abnormal trading volume in the current week is significantly driven by average abnormal trading volume in the last week. The respective regression model explains 14.6% of the variation in the dependent variable according to the adjusted  $R^2$ . If we add our investor attention proxy in column (2), the adjusted  $R^2$  significantly increases by 6.5 percentage points to 21.1%. Investor attention hence increases the predictive power of the regression model. High investor attention, as captured by our proxy based on Google Search volume for firm name and “stock” in local language, is associated with increased trading volume. A  $\beta_1 = 0.34$  indicates that if Google Search volume increases by 1% relative to median search volume of the last 8 weeks, average daily abnormal trading volume in the current week increases by 0.34% relative to median trading volume of the last 40 trading days.<sup>4</sup> This positive relation holds for the United States (5) and Europe (9), as well as for each European country (see Table 14 in the appendix)<sup>5</sup>. Therefore we can confirm that investor attention proxied by GSV for firm name and “stock” in local language is associated with increased trading

---

<sup>4</sup>Results are robust to calculating abnormal trading volume based on weekly trading volume relative to median of last 8 weeks’ trading volume.

<sup>5</sup>Except for Netherlands, where the coefficient is insignificant probably due to sample size.

Table 4: Trading volume

Dependent	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sample	$AVO_t$	$AVO_t$	$AVO_t$	$AVO_t$	$AVO_t$	$AVO_t$	$AVO_t$	$AVO_t$	$AVO_t$	$AVO_t$
	Full	Full	Full	US	US	US	US	Europe	Europe	Europe
$AVO_{t-1}$	0.376*** (0.0103)	0.330*** (0.00903)	0.367*** (0.00945)	0.351*** (0.0113)	0.305*** (0.00865)	0.345*** (0.00921)	0.322*** (0.0119)	0.399*** (0.0173)	0.352*** (0.0156)	0.386*** (0.0165)
$ASV_t^{Stock}$		0.340*** (0.0151)			0.334*** (0.0156)				0.349*** (0.0300)	
$ASV_t^{Name}$			0.454*** (0.0491)			0.342*** (0.0623)				0.675*** (0.0798)
$ASV_t^{Ticker}$							0.428*** (0.0496)			
Constant	-0.528*** (0.0792)	-0.591*** (0.0727)	-0.517*** (0.0817)	-0.548*** (0.116)	-0.624*** (0.106)	-0.545*** (0.110)	-0.580*** (0.1000)	-0.635*** (0.115)	-0.700*** (0.108)	-0.604*** (0.138)
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	45,353	45,353	45,353	26,884	26,884	26,884	26,173	18,469	18,469	18,469
Number of firms	210	210	210	119	119	119	117	91	91	91
(Adjusted) R <sup>2</sup>	0.146	0.211	0.170	0.132	0.207	0.150	0.189	0.162	0.216	0.197
Model comparison		(2) vs (3)		(5) vs (6)	(5) vs (7)					(9) vs (10)
Delta R <sup>2</sup>		0.041		0.057	0.018					0.019
Vuong Z-test		16.78***		17.09***	3.87***					5.20***
AIC	34,337	30,767	33,043	15,768	13,319	15,196	13,581	17,763	16,532	16,973
BIC	34,467	30,907	33,183	15,883	13,442	15,319	13,704	17,880	16,657	17,098

Firm-clustered standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Notes: Firm fixed-effect regression with clustered standard errors is applied.  $AVO_t$  is logarithm of average daily abnormal trading volume of week  $t$  calculated by dividing daily trading volume by median of last 40 days' trading volume,  $AVO_{t-1}$  is lagged  $AVO$  by one week,  $ASV_t^{Stock}$  is logarithm of abnormal Google Search volume for firm name and "stock" in local language divided by median of last 8 week's Google Search volume,  $ASV_t^{Name}$  is logarithm of abnormal Google Search volume for firm name divided by median of last 8 week's Google Search volume,  $ASV_t^{Ticker}$  is logarithm of abnormal Google Search volume for US ticker symbol divided by median of last 8 week's Google Search volume. Control variables include year dummies,  $Volat_t$  (historic volatility),  $Size_t$  (log market cap),  $BTM_t$  (book-to-market),  $InstOwn_t$  (institutional ownership),  $BASpread_t$  (bid-ask spread).



volume (hypothesis 1a).

In addition we compare our attention proxy consisting of firm name and “stock” in local language to other search term methodologies. Alternative search terms that have been proposed so far include raw firm name as well as ticker symbol in a US context. For the full sample, adjusted  $R^2$  in column (3) using ASV based on raw firm name is 17.0%. This is 4.1 percentage points lower than the adjusted  $R^2$  of regression model (2) using name and “stock” in local language. Hence, the regression model using ASV based on firm name and “stock” in local language is able to explain 4.1% variance in abnormal trading volume more than the regression model based on firm name. For the US sample, we additionally compare search volume for ticker symbol. Regression model (7) based on ticker symbol yields an  $R^2$  of 18.9%.<sup>6</sup> This is higher than for regression model (6) based on ASV for raw firm name with 15.0%. But compared to our proxy in model (5) both alternative search terms are less predictive for the US sample. If we take a closer look at Europe, regression model (10) based on ASV for firm name yields an  $R^2$  of 19.7%, which is lower than the  $R^2$  of model (9) based on our proxy at 21.6%. To test if those differences in  $R^2$  are significant, we conduct a formal test based on Vuong (1989). The Vuong method tests the Null that both competing non-nested models are equally distant from the true data generating process against the alternative hypothesis that one model is closer. Therefore, a significant Z score indicates that one model is superior to the other. The respective Vuong tests reveal that models based on  $ASV_t^{Stock}$  are significantly closer to the true model than the models based on alternative search terms. Hence, the differences in  $R^2$  are significantly different from zero. Obviously, also the information criteria AIC and BIC, which, like the Vuong test, are based on Loglikelihood (LL), are lower for the respective models using our search term approach, also indicating better model fit.<sup>7</sup>

It is important to emphasize that firm name and also ticker symbol is searched more frequently than firm name and “stock” in local language. Hence, more firm weeks with GSV scores different from 0 are available for these two proxies, than shown in Table 4. Therefore, we conduct additional regressions for which we do not limit the sample to firm-weeks with available  $ASV_t^{Stock}$  data. Results are presented in table 15 (in the appendix).

---

<sup>6</sup>Even if we exclude noisy ticker symbols (i.e., symbols with less than 3 characters), which gives us 20,786 firm week observations, the  $R^2$  is a little higher at 20.5%.

<sup>7</sup>AIC and BIC are calculated based on Loglikelihood (LL). Minimizing sum of squared residuals is equivalent to maximizing LL. Therefore, lower AIC and BIC values indicate better model fit/lower distance from the true model, as  $AIC = 2k - 2LL$  and  $BIC = k \times \ln(N) - 2LL$ , with k number of estimated parameters and N number of observations.

For the full sample, we gather 562,423 firm weeks based on Google Search volume for firm name.  $ASV_t^{Name}$  is significantly associated with abnormal trading volume and increases the  $R^2$  when entering the regression. But with  $R^2$  of 13% for the full sample, the predictive power is much lower. This again shows that firm name is a more noisy proxy for investor attention. If we limit the sample based on availability of  $ASV_t^{Stock}$ , the predictive power of  $ASV_t^{Name}$  increases. But this can potentially be explained, that in times of high investor attention, firm name partly reflects search behavior of attentive investors.

**Robustness check** So far we have shown that high attention in the current week leads to an increase in average abnormal trading volume in that same week. The skeptical reader could argue that it was not attention leading to trading volume, but that high trading volume rather leads to attention. It could well be that investors observe high trading volumes, which made them search these highly traded securities on Google.

To cope with this issue, we include time lags in our analysis and conduct a formal joint F-test for these time lagged variables. Results of the analysis are presented in Table 5. Column (1), (3) and (5) represent regressions with abnormal trading volume as dependent variable for the full, US and European sample respectively. Column (2), (4) and (6) are regressions with abnormal search volume as the dependent variable. Trading volume is significantly driven by lagged abnormal search volume. If search volume in the last week was high, trading volume is significantly higher in this week. The joint F-test for all four lags of abnormal search volume is statistically significant with  $F=26.69$ . Hence, lagged abnormal search volume leads to trading volume. We again can confirm that trading volume leads to trading volume itself ( $F=296.27$ ).

If we now turn to ASV as dependent variable we find that search volume also follows high trading volume. This relation is intuitive, because trading volume is obviously a stimulus, creating attention. We can conclude that lagged attention measured by abnormal Google Search Volume leads to trading volume, and at the same time trading volume leads to Google Search Volume measured attention. This is consistent with the findings of Dimpfl and Jank (2015), who show similar relationships in the context of Dow Jones index trading volume.

Table 5: Trading volume with time lags

Dependent	(1) $AVO_t$	(2) $ASV_t^{Stock}$	(3) $AVO_t$	(4) $ASV_t^{Stock}$	(5) $AVO_t$	(6) $ASV_t^{Stock}$
Sample	Full	Full	US	US	EU	US
$AVO_{t-1}$	0.363*** (0.0118)	0.0306*** (0.00516)	0.322*** (0.0122)	0.00936 (0.00714)	0.394*** (0.0189)	0.0487*** (0.00674)
$AVO_{t-2}$	-0.0197** (0.00929)	-0.0301*** (0.00490)	-0.0371*** (0.00966)	-0.0376*** (0.00736)	-0.00877 (0.0158)	-0.0262*** (0.00633)
$AVO_{t-3}$	-0.0209*** (0.00611)	0.0127*** (0.00441)	-0.0296*** (0.00750)	0.0168** (0.00657)	-0.0127 (0.00976)	0.00790 (0.00592)
$AVO_{t-4}$	-0.0230*** (0.00703)	-0.000694 (0.00438)	0.0234** (0.0102)	0.00778 (0.00770)	-0.0536*** (0.00835)	-0.00517 (0.00519)
$ASV_{t-1}^{Stock}$	0.0813*** (0.00903)	0.403*** (0.0107)	0.0870*** (0.0105)	0.396*** (0.0146)	0.0770*** (0.0144)	0.411*** (0.0157)
$ASV_{t-2}^{Stock}$	0.00432 (0.00953)	0.0828*** (0.00754)	0.0256** (0.0105)	0.0774*** (0.00934)	-0.0179 (0.0161)	0.0906*** (0.0121)
$ASV_{t-3}^{Stock}$	0.0156 (0.00962)	-0.00263 (0.00783)	0.0133 (0.0119)	-0.00299 (0.0105)	0.0222 (0.0158)	-0.00112 (0.0119)
$ASV_{t-4}^{Stock}$	-0.00715 (0.00773)	-0.0679*** (0.00714)	-0.00755 (0.00949)	-0.0729*** (0.0106)	-0.0123 (0.0129)	-0.0638*** (0.00947)
Constant	-0.513*** (0.107)	0.0650 (0.0491)	-0.676*** (0.111)	0.152* (0.0798)	-0.615*** (0.176)	0.0752 (0.0706)
Controls	yes	yes	yes	yes	yes	yes
Observations	33,969	33,969	18,330	18,330	15,639	15,639
Number of firms	0.151	0.208	0.133	0.194	0.173	0.227
(Adjusted) R <sup>2</sup>	209	209	119	119	90	90
F-test: $AVO_{t-1/4} = 0$	296.27***	16.22***	211.6***	6.79***	176.27***	16.53***
F-test: $ASV_{t-1/4}^{Stock} = 0$	26.69***	418.72***	37.82***	241.02***	8.06***	203.54***

Firm-clustered standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Notes: Firm fixed-effect regression with clustered standard errors is applied.  $AVO_t$  is logarithm of average daily abnormal trading volume of week t calculated by dividing daily trading volume by median of last 40 days' trading volume,  $AVO_{t-j}$  is AVO lagged by j weeks,  $ASV_t^{Stock}$  is logarithm of abnormal Google Search volume for firm name and "stock" in local language divided by median of last 8 week's Google Search volume,  $ASV_{t-j}^{Stock}$  is ASV lagged by j weeks. F-test is joint F-test with Null that AVO or ASV respectively with j=1,2,3 and 4 lags are jointly equal to 0. Control variables include year dummies,  $Vola_t$  (historic volatility),  $Size_t$  (log market cap),  $BTM_t$  (book-to-market),  $InstOwn_t$  (institutional ownership),  $BASpread_t$  (bid-ask spread).

We have shown that searches in the last week lead to more trading in the current week. It is, however, hard to believe, that the majority of investors wait up to one week after searching a stock before trading (if we abstract from searches made on Sunday leading to trading on Monday).

Therefore, in order to further investigate this vicious circle of attention leading to trading and trading leading to attention, we make use of the daily Google Search volume data. Table 6 shows regressions explaining abnormal trading volume based on daily Google Search volume. For all samples, searches yesterday lead to higher trading volume today. This provides further robustness that GSV predicts trading volume when using

lagged scores and at least explains trading volume when using simultaneous scores.

Table 6: Trading volume (daily panel)

Dependent	(1) $AVO_t$	(2) $AVO_t$	(3) $AVO_t$	(4) $AVO_t$	(5) $AVO_t$	(6) $AVO_t$
Sample	Full	Full	US	US	Europe	Europe
$AVO_{t-1}$	0.530*** (0.0263)	0.502*** (0.0254)	0.585*** (0.0158)	0.557*** (0.0135)	0.456*** (0.0485)	0.430*** (0.0471)
$ASV_{t-1}^{Stock}$		0.0942*** (0.00988)		0.0791*** (0.00752)		0.119*** (0.0251)
Constant	0.169 (0.136)	0.171 (0.140)	0.257** (0.0996)	0.259** (0.0990)	-0.129 (0.119)	-0.107 (0.125)
Controls	yes	yes	yes	yes	yes	yes
Observations	53,644	53,644	36,939	36,939	16,705	16,705
Number of firms	100	100	64	64	36	36
(Adjusted) $R^2$	0.286	0.296	0.348	0.357	0.218	0.229

Firm-clustered standard errors in parentheses

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Notes: Firm fixed-effect regression with clustered standard errors is applied.  $AVO_t$  is logarithm of abnormal trading volume of day  $t$  calculated by dividing average trading volume per week by median of last 40 days' trading volume,  $AVO_{t-1}$  is lagged AVO by one day,  $ASV_t^{Stock}$  is logarithm of abnormal Google Search volume for firm name and "stock" in local language divided by median of last 56 days' Google Search volume. Control variables include year and day of week dummies,  $Volat_t$  (historic volatility),  $Size_t$  (log market cap),  $BTM_t$  (book-to-market),  $InstOwn_t$  (institutional ownership),  $BASpread_t$  (bid-ask spread).

### 5.1.2 Return

Investor attention is associated with increased trading volume. Next we investigate, whether those trades are purchases or sales of assets. Results of this stock return analysis are presented in Table 7. Column (1) confirms hypothesis 1b that if the market is in fear, returns are negative. As expected, column (2) reveals that attention alone does not explain any returns.  $R^2$  of the regression model is almost 0 and the beta coefficient for  $ASV_t^{Stock}$  is insignificant.

But when we add the market fear dummy, investor attention can be used to explain returns. Column (3) shows that attention has a positive impact on return, if the level of the market volatility index decreases or stays constant, i.e.,  $Fear_t = 0$ . This is in accordance with Barber and Odean (2008), who find that retail investors are "net buyers" of attention grabbing stocks. As retail investors typically do not sell short, they are restricted to sell the few stocks they already hold in their portfolio.

Table 7: Return and investor sentiment

Dependent	(1)	(2)	(3)	(4)	(5)
	$r_t$	$r_t$	$r_t$	$r_t$	$r_t$
Sample	Full	Full	Full	US	Europe
$Fear_t$	-0.0265*** (0.000947)		-0.0262*** (0.000922)	-0.0254*** (0.00108)	-0.0273*** (0.00162)
$ASV_t^{Stock}$		-0.00192 (0.00165)	0.0111*** (0.00246)	0.0129*** (0.00353)	0.00840*** (0.00311)
$Fear_t \times ASV_t^{Stock}$			-0.0240*** (0.00490)	-0.0279*** (0.00716)	-0.0179*** (0.00552)
Constant	-0.0316** (0.0128)	-0.0506*** (0.0125)	-0.0304** (0.0129)	-0.0178 (0.0117)	-0.0754*** (0.0206)
Controls	yes	yes	yes	yes	yes
Observations	45,353	45,353	45,353	26,884	18,469
Number of firms	210	210	210	119	91
(Adjusted) R <sup>2</sup>	0.063	0.011	0.067	0.073	0.062

Firm-clustered standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Notes: Firm fixed-effect regression with clustered standard errors is applied.  $r_t$  is raw return of week t,  $Fear_t$  is dummy taking the value of 1 if the level of the corresponding market volatility index in week t increases,  $ASV_t^{Stock}$  is logarithm of abnormal Google Search volume for firm name and “stock” in local language divided by median of last 8 week’s Google Search volume. Control variables include year dummies,  $Vola_t$  (historic volatility),  $Size_t$  (log market cap),  $BTM_t$  (book-to-market),  $InstOwn_t$  (institutional ownership),  $BASpread_t$  (bid-ask spread).

Hence, on average attention leads to buying rather than selling. Additionally we find that if the market is in fear, high attention leads to negative returns instead. In turbulent times, investors are likely to be negatively influenced by uncertainty and turmoil. Therefore, their attention on average leads to selling or negative returns. Results are robust for the US as well as the European sub sample (see column (4) and (5)).

**Robustness check** If we again make use of our daily sample to include time lags, search volume of yesterday predicts positive returns today for  $Fear_t = 0$  and negative returns for  $Fear_t = 1$ , however, only significant for the full sample and the US sub sample.<sup>8</sup>

## 5.2 Market efficiency

Investor attention, as captured by our novel attention proxy, is associated with trading activity. It explains trading volume and after controlling for sentiment even market

<sup>8</sup>Results not reported for parsimony reasons, but available upon request.

returns. Now we assess, whether this trading activity leads to more efficiency, or, more precisely, whether retail investor attention can enhance market efficiency.

### 5.2.1 Incorporation of idiosyncratic information

A security is priced efficiently if its price is driven by firm-specific or idiosyncratic information, rather than being based on market wide-factors. Following Bris et al. (2007) and Chang et al. (2014) we assess the impact of attention on pricing efficiency using the market-model  $R^2$  approach. We regress the return of a security on the return of the market and calculate the  $R^2$  of that regression. A high  $R^2$  implies low pricing efficiency, because variation in the stock price can to a large extent be explained by variation in the market return. In that logic, low  $R^2$  implies higher pricing efficiency, as returns are based on idiosyncratic or firm-specific information. Results are shown in Table 8.

For the full sample, the average  $R^2$  of the 206 analyzed firms in weeks with high attention is 42.0%, in weeks with low attention the  $R^2$  is 46.4%. Hence, it is 4.4 percentage points lower in weeks with high attention (statistically significant at 1 % level). This suggests that high retail investor attention is associated with increased pricing efficiency.

We assume that retail investors tend to base trading decisions on sentiment, i.e., believes rather than information. Therefore, we further distinguish the state of the market (bullish and bearish) as well as market fear (unfearful and fearful), as we expect different behavior for these different states. If we split the sample into two groups based on the sign of market return, an additional conclusion can be drawn. For positive market returns (bullish), the efficiency enhancing effect is even more pronounced. While for weeks with low attention, on average 35.8% of the variance in the return of the security can be explained by variance in the return of the market, for weeks with high attention only 27.7% can be explained. This yields an even higher delta or efficiency gain of -8.1 percentage points in terms of  $R^2$  difference.

For negative market returns (bearish), the effect is the opposite, i.e.,  $R^2$  Bearish is significantly higher for firm weeks with high attention, implying reduced market efficiency. But the delta of +5.8 percentage points efficiency loss in bearish markets is lower than the efficiency gain in bullish markets. Therefore, the overall effect still leads to a negative delta and hence to an enhancement of market efficiency. This asymmetry of investor behavior between bullish and bearish markets can be explained, as retail investors are typically restricted to sell securities they already hold in their portfolio and usually do not sell short (similar to Barber and Odean, 2008). Therefore, the effect in bearish markets, in which selling could cause stronger co-movements of returns, is weaker.

Table 8: Market Model  $R^2$ 

	N	High	Low	Delta	t-test
<b>Full</b>					
$R^2$ Overall	206	42.0%	46.4%	-4.4%	-4.38***
$R^2$ Bullish	156	27.7%	35.8%	-8.1%	-5.90***
$R^2$ Bearish	161	41.0%	35.2%	5.8%	3.72***
$R^2$ Unfearful	172	37.2%	44.1%	-6.9%	-4.79***
$R^2$ Fearful	181	47.4%	47.5%	0.0%	-0.03
<b>Europe</b>					
$R^2$ Overall	89	44.3%	49.2%	-4.8%	-2.8***
$R^2$ Bullish	59	28.5%	38.4%	-9.9%	-4.16***
$R^2$ Bearish	66	45.7%	40.4%	5.4%	2.10**
$R^2$ Unfearful	68	41.8%	49.4%	-7.7%	-2.84***
$R^2$ Fearful	76	52.1%	53.0%	-0.9%	-0.45
<b>US</b>					
$R^2$ Overall	117	40.3%	44.3%	-4.0%	-3.44***
$R^2$ Bullish	97	27.2%	34.2%	-7.0%	-4.19***
$R^2$ Bearish	95	37.8%	31.7%	6.1%	3.10***
$R^2$ Unfearful	104	34.3%	40.6%	-6.3%	-3.99***
$R^2$ Fearful	105	44.1%	43.5%	0.6%	0.31

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Notes:  $R^2$  values are estimated for each firm using ordinary least square (OLS) regression model with  $r_{it} = \alpha_i + \beta_{1i}mkt_t + \epsilon_{it}$ . Thereby, N is number of regressions/firms, High is average of all regression  $R^2$  estimated based on firm weeks with high abnormal attention, Low is average of all regression  $R^2$  based on firm weeks with low abnormal attention, Delta is average difference in  $R^2$  between High and Low weeks' regressions and t-test is statistical significance of Delta. Bullish is based on firm weeks with positive market returns, Bearish is based on firm weeks with negative market returns, Unfearful is based on firm weeks with negative or stable changes in the level of the corresponding market volatility index, fearful is based on firm weeks with positive changes in the level of the corresponding market volatility index.

A possible explanation for this deterioration in efficiency could be that in times of bearish markets, retail investors could be caught by the herding effect (Barber et al., 2009). Investors observe a decline in the market and are attentive towards stocks in their portfolio by searching the internet for stock related information. Then, they tend to sell those stocks due to the negative market wide factors and behavior of other investors, irrespective of any firm-specific information. This behavior leads to higher co-movement with market returns and hence to the observed deterioration in pricing efficiency.

In addition, we split the sample into “unfearful” and “fearful” weeks. “Unfearful” weeks are those, in which the level of the market volatility index does not change or decreases, implying low financial uncertainty. For the “unfearful” weeks, high attention increases efficiency, the  $R^2$  is 6.9 percentage points lower in weeks with high attention compared to those with low attention. In fearful weeks, no significant difference between the high and low group can be observed. If financial uncertainty is high, retail investors’ behavior does not cause any changes in the level of pricing efficiency. Either, because investors do not act at all, or, because their actions at least do not lead to stronger co-movement with the market.

All effects can be observed for the European as well as the US sub sample. On a per country level for the European firms results are similar, but partly insignificant due to the low number of firms per country (see Table 17 in the appendix).

**Robustness check** Results are robust to using Google search volume for firm name (see Table 16 in the appendix). Nevertheless, using firm names leads to less pronounced results. The difference between high and low is smaller and the delta for bearish markets is insignificant. This again shows that firm name is a more noisy proxy for investor attention. Furthermore, results are robust to using “symmetric” allocation<sup>9</sup> of the high and low group. Also, using return of the current and past week instead of current week only in order to identify a bullish or bearish market produces robust results. However, bullish and bearish deltas are only significant for the United States sub sample. For the full sample and for Europe, the deltas in the bearish market are then insignificant.

### 5.2.2 Return predictability

In order to stress robustness of our findings we apply an additional efficiency measure. We test, whether attention can mitigate a market anomaly, which imposes severe violations to market efficiency. If a market is (weak form) efficient, all historic information should be reflected in asset prices. Returns should follow a pure random walk. Hence, historic stock returns cannot be applied in order to predict future returns. To test whether there is any return predictability in our sample and if it is mitigated by investor attention, we conduct an analysis as presented in Table 9.

For the full sample shown in column (1), return lagged by one week  $r_{t-1}$  indeed negatively predicts return in the current week. From our previous analysis we conclude

---

<sup>9</sup>Allocation relative to the median of all abnormal search volumes per firm, i.e., high for  $ASV > Median(ASV)$  and low for  $ASV < Median(ASV)$



Table 9: Return predictability

	(1)	(2)	(3)
Dependent	$r_t$	$r_t$	$r_t$
Sample	Full	US	Europe
$r_{t-1}$	-0.0561*** (0.0112)	-0.0504*** (0.0174)	-0.0639*** (0.0165)
$ASV_{t-1}$	-0.00926*** (0.00130)	-0.00895*** (0.00148)	-0.00902*** (0.00260)
$r_{t-1} \times ASV_{t-1} \times Bullish_{t-1}$	0.0644*** (0.0199)	0.0705*** (0.0142)	0.0286 (0.0854)
$r_{t-1} \times ASV_{t-1} \times Bearish_{t-1}$	-0.178*** (0.0449)	-0.190*** (0.0575)	-0.163** (0.0716)
Constant	-0.0653*** (0.0136)	-0.0589*** (0.0157)	-0.0958*** (0.0215)
Controls	yes	yes	yes
Observations	53,914	33,327	20,587
Number of firms	216	122	94
(Adjusted) R <sup>2</sup>	0.021	0.022	0.022

Firm-clustered standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Notes: Firm fixed-effect regression with clustered standard errors is applied.  $r_t$  is raw return in week t,  $r_{t-1}$  is lagged return by one week,  $ASV_{t-1}$  is logarithm of abnormal Google Search volume for firm name and “stock” in local language divided by median of last 8 week’s Google Search Volume lagged by one week,  $r_{t-1} \times ASV_{t-1} \times Bullish_{t-1}$  is the interaction term of raw return and Google search volume lagged by one week if market return lagged by one week is greater or equal to zero and zero otherwise,  $r_{t-1} \times ASV_{t-1} \times Bearish_{t-1}$  is the interaction term of raw return and Google search volume lagged by one week if market return lagged by one week is smaller than zero and zero otherwise. Control variables include year dummies,  $Vola_t$  (historic volatility),  $Size_t$  (log market cap),  $BTM_t$  (book-to-market),  $InstOwn_t$  (institutional ownership),  $BASpread_t$  (bid-ask spread).

that in times of bullish markets, attention enhances efficiency and in times of bearish markets, attention deteriorates efficiency. Therefore, we distinguish the state of the market for this analysis as well. We introduce a dummy variable  $Bullish_{t-1}$  ( $Bearish_{t-1}$ ) taking the value of 1 if the market return in the last week was positive (negative) and 0 otherwise.

The interaction effect of last week’s return with last week’s Google search volume in a bullish market  $r_{t-1} \times ASV_{t-1} \times Bullish_{t-1}$  has a positive impact on the return in the current week. This effect runs against the negative return prediction effect of  $r_{t-1}$ . If investor attention in the last week was high and the market increased, available

information was better and more efficiently incorporated into asset prices. Hence, historic information can be less used to predict asset prices in the current week and therefore return predictability is less pronounced. This implies, the market is more efficient, if investor attention was high (H2b).

For bearish markets, however, the effect is the opposite:  $r_{t-1} \times ASV_{t-1} \times Bearish_{t-1}$  has a negative impact on the return in the current week, which leads to an even stronger return predictability effect. This supports our previous findings that in times of bearish markets, retail investors tend to act irrationally, leading to less efficient markets or more pronounced return predictability.

Both interaction terms are significant in the full sample (1) and the US sub sample (2). In Europe, signs are consistent, but insignificant for bullish markets. If we further look into each European country (see Table 18 in the appendix), return predictability for one week lagged returns is significantly present in France, Germany, Italy and Switzerland. The extenuation of return predictability through attention (for bullish markets) can significantly be observed in France and Switzerland. For the other countries, signs are consistent but coefficients are insignificant.

## 6 Conclusion

In this study we propose a novel investor attention proxy based on Google Search volume. We use a modified search term consisting of firm name and “stock” in local language and verify that attention as captured by our proxy explains trading activity. If we control for market fear measured by changes in the level of the market volatility index, investor attention explains stock returns. If market fear is low, attention is associated with positive returns, if fear is high, attention is associated with negative returns. We then show that high retail investor attention is associated with a higher degree of market efficiency. Attention leads to better incorporation of idiosyncratic information, which we interpret as increased pricing efficiency. Security prices are then to a larger extent based on firm-specific information rather than being driven by market-wide factors. This effect, however, deviates depending on the state of the market. Retail investor attention enhances pricing efficiency overall with this effect being even more pronounced in times of bullish markets. But in bearish markets the opposite is true: attention has a deteriorating impact on efficiency. We explain these findings with irrational herding behavior of retail investors in declining markets. We further stress robustness of our findings and show that retail investor attention mitigates return predictability in bullish markets.

## References

- Aouadi, A., Arouri, M., and Teulon, F. (2013). Investor attention and stock market activity: Evidence from France. *Economic Modelling*, 35:674–681.
- Baker, M. and Wurgler, J. (2007). Investor Sentiment in the Stock Market. *The Journal of Economic Perspectives*, 21(2):129–151.
- Bank, M., Larch, M., and Peter, G. (2011). Google search volume and its influence on liquidity and returns of German stocks. *Financial Markets and Portfolio Management*, 25(3):239–264.
- Barber, B. M. and Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies*, 21(2):785–818.
- Barber, B. M., Odean, T., and Zhu, N. (2009). Do Retail Trades Move Markets? *Review of Financial Studies*, 22(1):151–186.
- Bris, A., Goetzmann, W. N., and Zhu, N. (2007). Efficiency and the Bear: Short Sales and Markets Around the World. *The Journal of Finance*, 62(3):1029–1079.
- Chang, E. C., Luo, Y., and Ren, J. (2014). Short-selling, margin-trading, and price efficiency: Evidence from the Chinese market. *Journal of Banking & Finance*, 48:411–424.
- Da, Z., Engelberg, J., and Gao, P. (2011). In Search of Attention. *The Journal of Finance*, 66(5):1461–1499.
- Da, Z., Engelberg, J., and Gao, P. (2015). The Sum of All FEARS Investor Sentiment and Asset Prices. *Review of Financial Studies*, 28(1):1–32.
- Daniel, K., Hirshleifer, D., and Teoh, S. H. (2002). Investor psychology in capital markets: Evidence and policy implications. *Journal of Monetary Economics*, 49(1):139–209.
- De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy*, 98(4):703–738.

- Dellavigna, S. and Pollet, J. M. (2009). Investor Inattention and Friday Earnings Announcements. *The Journal of Finance*, 64(2):709–749.
- Dimpfl, T. and Jank, S. (2015). Can Internet Search Queries Help to Predict Stock Market Volatility? *European Financial Management*, 9999:1–22.
- Doukas, J. A., Ball, R., Daniel, K., French, K., Ross, S., and Shanken, J. (2002). Rationality of Capital Markets. *European Financial Management*, 8(2):229–247.
- Drake, M. S., Roulstone, D. T., and Thornock, J. R. (2012). Investor Information Demand: Evidence from Google Searches Around Earnings Announcements. *Journal of Accounting Research*, 50(4):1001–1040.
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2):383–417.
- Fama, E. F. (1991). Efficient Capital Markets: II. *The Journal of Finance*, 46(5):1575–1617.
- Fang, L. and Peress, J. (2009). Media Coverage and the Cross-section of Stock Returns. *The Journal of Finance*, 64(5):2023–2052.
- Fricke, E., Fung, S., and Goktan, M. S. (2014). Google Search, Information Uncertainty, and Post-Earnings Announcement Drift. *Journal of Accounting and Finance*, 14(2):11–27.
- Grinblatt, M. and Keloharju, M. (2000). The investment behavior and performance of various investor types: a study of Finland’s unique data set. *Journal of Financial Economics*, 55(1):43–67.
- Hirshleifer, D., Lim, S. S., and Teoh, S. H. (2009). Driven to Distraction: Extraneous Events and Underreaction to Earnings News. *The Journal of Finance*, 64(5):2289–2325.
- Huberman, G. and Regev, T. (2001). Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar. *The Journal of Finance*, 56(1):387–396.
- Jacobs, H. and Weber, M. (2011). The Trading Volume Impact of Local Bias: Evidence from a Natural Experiment. *Review of Finance*, 0:1–35.

- Malkiel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. *The Journal of Economic Perspectives*, 17(1):59–82.
- Merton, R. C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, 42(3):483–510.
- Odean, T. (1998). Are Investors Reluctant to Realize Their Losses? *The Journal of Finance*, 53(5):1775–1798.
- Siganos, A. (2013). Google attention and target price run ups. *International Review of Financial Analysis*, 29:219–226.
- Vozlyublennaiia, N. (2014). Investor attention, index performance, and return predictability. *Journal of Banking & Finance*, 41:17–35.
- Vuong, Q. H. (1989). Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica*, 57(2):307–333.
- Whaley, R. E. (2000). The Investor Fear Gauge. *The Journal of Portfolio Management*, 26(3):12–17.

# Appendix

Table 10: Market volatility indexes

Country	Volatility Index	Statistics			Quartiles					Pearson Correlation			
		N	Mean	St.Dev.	Min	0.25	Med	0.75	Max	(1)	(2)	(3)	(4)
(1) Euro countries	VSTOXX Volatility Index	2,909	23.13	9.37	11.60	16.82	20.88	26.27	87.51	1.00			
(2) Switzerland	VSMI Volatility index	2,909	18.09	8.10	9.24	13.36	15.67	19.89	84.90	0.93	1.00		
(3) United Kingdom	FTSE 100 Volatility Index	2,909	18.96	8.69	9.10	13.06	16.25	22.05	75.54	0.96	0.97	1.00	
(4) United States	CBOE SPX Volatility VIX	2,909	19.56	9.53	9.89	13.42	16.60	22.37	80.86	0.96	0.95	0.97	1.00

Notes: Sample consists of 2,909 trading days between January 2004 and February 2015.

Table 11: Variable description

Symbol	Description	Calculation	Source
$ASV_t^{Stock}$	Abnormal Search Volume "Stock"	Logarithm of GSV for firm name and "stock" in local language relative to median of last 8 weeks' (56 days') GSV	Google Trends
$ASV_t^{Name}$	Abnormal Search Volume "Name"	Logarithm of GSV for firm name relative to median of last 8 weeks' (56 days') GSV	Google Trends
$ASV_t^{Ticker}$	Abnormal Search Volume "Ticker"	Logarithm of GSV for ticker symbol relative to median of last 8 weeks' (56 days') GSV	Google Trends
$AVO_t$	Abnormal trading volume	Logarithm of trading volume relative to median of last 40 days' trading volume (average for weekly data)	Thomson
$BASpread_t$	Bid-ask spread	Ask price minus low price deflated by mid-point (average for weekly data)	Thomson
$BTM_t$	Book-to-Market	Book value of equity divided by total market capitalization divided by 1,000 (average for weekly data)	Thomson
$Fear_t$	Investor sentiment	Dummy taking value of 1, if level of corresponding market volatility index increases	Thomson
$InstOwn_t$	Institutional Ownership	Number of shares held by institutional owners divided by shares outstanding	Thomson
$mkt_t$	Market return	Return of the corresponding market portfolio	Thomson
$r_t$	Raw return	Logarithm return calculated based on total return index (RI)	Thomson
$Size_t$	Size	Logarithm of total market capitalization (average for weekly data)	Thomson
$Volat_t$	Historic volatility	Standard deviation of last 180 trading days (average for weekly data)	Thomson

Table 12: Summary statistics (weekly data)

Variable	Statistics			Pearson correlation											
	N	Mean	St.Dev.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) $AVO_t$	45,353	-1.5695	0.3853	1.0000											
(2) $AVO_{t-1}$	45,353	-1.5607	0.3833	0.3565	1.0000										
(3) $ASV_t^{Stock}$	45,353	0.0108	0.2916	0.3242	0.1604	1.0000									
(4) $ASV_t^{Name}$	45,353	0.0043	0.1319	0.1442	0.0413	0.3283	1.0000								
(5) $ASV_t^{Ticker}$	26,173	0.0084	0.1979	0.2787	0.1133	0.2700	0.1914	1.0000							
(6) $r_t$	45,353	0.0011	0.0579	-0.0647	-0.0299	-0.0140	-0.0014	0.0042	1.0000						
(7) $Fear_t$	45,353	0.4634	0.4987	0.1323	0.0026	0.0258	0.0005	0.0021	-0.2416	1.0000					
(8) $Volat_{it}$	45,353	0.0217	0.0138	-0.0302	-0.0126	-0.0289	-0.0066	-0.0218	0.0227	-0.0297	1.0000				
(9) $Size_{it}$	45,353	10.1613	1.4305	-0.0214	-0.0224	0.0014	-0.0211	0.0244	0.0016	-0.0014	-0.3400	1.0000			
(10) $BTM_{it}$	45,353	0.8838	2.2604	0.0127	0.0131	-0.0015	-0.0075	-0.0216	-0.0060	-0.0059	0.3195	-0.1013	1.0000		
(11) $InstOwn_{it}$	45,353	5.8779	8.6812	0.0150	0.0138	0.0053	0.0223	-0.0045	-0.0041	0.0040	0.0639	-0.3572	-0.0769	1.0000	
(12) $BASpread_{it}$	45,353	0.0019	0.0081	0.0014	0.0022	0.0308	0.0232	-0.0002	-0.0202	0.0042	0.0835	-0.0489	0.0270	0.0141	1.0000



Table 13: Summary statistics (daily data)

Variable	Statistics			Pearson correlation									
	N	Mean	St.Dev.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) $AVO_t$	53,644	0.0308	0.4628	1.0000									
(2) $AVO_{t-1}$	53,644	0.0316	0.4618	0.5284	1.0000								
(3) $ASV_t^{Stock}$	53,644	0.1239	0.5040	0.2975	0.2570	1.0000							
(4) $r_t$	53,644	0.0004	0.0252	-0.0056	0.0063	0.0124	1.0000						
(5) $Fear_t$	53,644	0.4607	0.4985	0.0337	-0.0109	-0.0101	-0.2862	1.0000					
(6) $Volat_{it}$	53,644	0.0207	0.0159	0.0310	0.0366	0.0315	-0.0056	-0.0200	1.0000				
(7) $Size_{it}$	53,644	10.6865	1.3233	-0.0201	-0.0197	0.0136	0.0154	-0.0043	-0.3052	1.0000			
(8) $BTM_{it}$	53,644	0.6964	1.2699	0.0127	0.0114	0.0361	-0.0461	0.0012	0.2868	-0.1586	1.0000		
(9) $InstOwn_{it}$	53,644	5.3555	6.9800	0.0048	0.0044	-0.0070	0.0095	0.0020	-0.0322	-0.2908	-0.1395	1.0000	
(10) $BASpread_{it}$	53,644	0.0015	0.0066	0.0298	0.0245	0.0256	0.0311	0.0097	0.2208	-0.1816	0.1308	-0.0030	1.0000

Table 14: Trading volume (per European country)

Dependent	(1) $AVO_t$	(2) $AVO_t$	(3) $AVO_t$	(4) $AVO_t$	(5) $AVO_t$	(6) $AVO_t$	(7) $AVO_t$	(8) $AVO_t$
Sample	Belgium	France	Germany	Italy	Netherlands	Spain	Switzerland	UK
$AVO_{t-1}$	0.350*** (0.00428)	0.384*** (0.0112)	0.264*** (0.0230)	0.498*** (0.0290)	0.297 (0.0623)	0.417*** (0.0597)	0.277** (0.0498)	0.279*** (0.0319)
$ASV_t^{Stock}$	0.324* (0.0265)	0.240*** (0.0345)	0.665*** (0.0459)	0.346*** (0.0371)	0.267* (0.0343)	0.352*** (0.0670)	0.319** (0.0532)	0.168** (0.0593)
Constant	-1.081* (0.145)	-0.288 (0.273)	-0.855** (0.314)	-1.501*** (0.164)	1.819 (1.112)	-0.0574 (0.880)	-0.453 (0.964)	-0.994*** (0.217)
Controls	yes	yes	yes	yes	yes	yes	yes	yes
Observations	306	5,914	4,322	1,420	456	1,254	585	4,212
Number of firms	2	22	30	9	2	6	3	17
(Adjusted) R <sup>2</sup>	0.262	0.218	0.321	0.346	0.251	0.357	0.240	0.108

Firm-clustered standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Notes: Firm fixed-effect regression with clustered standard errors is applied.  $AVO_t$  is logarithm of average daily abnormal trading volume of week t calculated by dividing daily trading volume by median of last 40 days' trading volume,  $AVO_{t-1}$  is lagged AVO by one week,  $ASV_t^{Stock}$  is logarithm of abnormal Google Search volume for firm name and "stock" in local language divided by median of last 8 week's Google Search volume,  $ASV_t^{Name}$  is logarithm of abnormal Google Search volume for firm name divided by median of last 8 week's Google Search volume,  $ASV_t^{Ticker}$  is logarithm of abnormal Google Search volume for US ticker symbol divided by median of last 8 week's Google Search volume. Control variables include year dummies,  $Vola_t$  (historic volatility),  $Size_t$  (log market cap),  $BTM_t$  (book-to-market),  $InstOwn_t$  (institutional ownership),  $BASpread_t$  (bid-ask spread).

Table 15: Trading volume (based on GSV for firm name and ticker)

Dependent	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$AVO_t$	$AVO_t$	$AVO_t$	$AVO_t$	$AVO_t$	$AVO_t$	$AVO_t$
Sample	Full	Full	US	US	US	Europe	Europe
$AVO_{t-1}$	0.339*** (0.00368)	0.333*** (0.00357)	0.354*** (0.00340)	0.348*** (0.00325)	0.344*** (0.00381)	0.337*** (0.00405)	0.331*** (0.00394)
$ASV_t^{Name}$		0.340*** (0.00950)		0.326*** (0.0177)			0.344*** (0.0111)
$ASV_t^{Ticker}$					0.289*** (0.0268)		
Constant	-0.569*** (0.0261)	-0.573*** (0.0261)	-0.720*** (0.0452)	-0.729*** (0.0432)	-0.759*** (0.0438)	-0.555*** (0.0300)	-0.558*** (0.0300)
Controls	yes	yes	yes	yes	yes	yes	yes
Observations	562,423	562,423	145,622	145,622	137,862	416,801	416,801
Number of firms	1,620	1,620	438	438	424	1,182	1,182
(Adjusted) R <sup>2</sup>	0.119	0.132	0.134	0.161	0.158	0.118	0.130
AIC	910,244	901,570	91,580	86,959	82,440	757,101	751,392
BIC	910,424	901,761	91,719	87,107	82,589	757,276	751,578

Firm-clustered standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Notes: Firm fixed-effect regression with clustered standard errors is applied.  $AVO_t$  is logarithm of average daily abnormal trading volume of week t calculated by dividing daily trading volume by median of last 40 days' trading volume,  $AVO_{t-1}$  is lagged AVO by one week,  $ASV_t^{Name}$  is logarithm of abnormal Google Search volume for firm name divided by median of last 8 week's Google Search volume,  $ASV_t^{Ticker}$  is logarithm of abnormal Google Search volume for US ticker symbol divided by median of last 8 week's Google Search volume. Control variables include year dummies,  $Volat_t$  (historic volatility),  $Size_t$  (log market cap),  $BTM_t$  (book-to-market),  $InstOwn_t$  (institutional ownership),  $BASpread_t$  (bid-ask spread).

Table 16: Market model  $R^2$  (based on GSV for firm name)

	<b>N</b>	<b>High</b>	<b>Low</b>	<b>Delta</b>	<b>t-test</b>
<b>Full</b>					
$R^2$ Overall	1496	29.5%	32.5%	-3.0%	-10.36***
$R^2$ Bullish	961	18.7%	23.1%	-4.4%	-10.45***
$R^2$ Bearish	1297	29.0%	28.9%	0.1%	0.20
$R^2$ Unfearful	1201	25.9%	30.0%	-4.1%	-10.54***
$R^2$ Fearful	1404	33.6%	35.5%	-1.9%	-4.44***
<b>Europe</b>					
$R^2$ Overall	1059	26.2%	28.8%	-2.6%	-7.87***
$R^2$ Bullish	561	16.9%	20.3%	-3.4%	-6.9***
$R^2$ Bearish	902	27.1%	27.3%	-0.1%	-0.21
$R^2$ Unfearful	782	24.5%	27.6%	-3.2%	-6.87***
$R^2$ Fearful	979	31.1%	32.9%	-1.8%	-3.73***
<b>US</b>					
$R^2$ Overall	437	37.5%	41.5%	-4.0%	-6.82***
$R^2$ Bullish	400	21.2%	27.0%	-5.8%	-7.91***
$R^2$ Bearish	395	33.4%	32.8%	0.6%	0.63
$R^2$ Unfearful	419	28.5%	34.4%	-5.9%	-8.31***
$R^2$ Fearful	425	39.3%	41.3%	-2.0%	-2.42**

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Notes:  $R^2$  values are estimated for each firm using ordinary least square (OLS) regression model with  $r_{it} = \alpha_i + \beta_{1i}mkt_t + \epsilon_{it}$ . Thereby, N is number of regressions/firms, High is average of all regression  $R^2$  estimated based on firm weeks with high abnormal attention, Low is average of all regression  $R^2$  based on firm weeks with low abnormal attention, Delta is average difference in  $R^2$  between High and Low weeks' regressions and t-test is statistical significance of Delta. Bullish is based on firm weeks with positive market returns, Bearish is based on firm weeks with negative market returns, Unfearful is based on firm weeks with negative or stable changes in the level of the corresponding market volatility index, fearful is based on firm weeks with positive changes in the level of the corresponding market volatility index.

Table 17: Market model  $R^2$  (per European country)

	N	High	Low	Delta	t-test
Belgium					
$R^2$ Overall	2	41.9%	37.7%	4.2%	0.53
$R^2$ Bullish	1	19.9%	41.9%	-22.0%	
$R^2$ Bearish	2	56.0%	26.1%	29.9%	7.67**
$R^2$ Unfearful	1	63.4%	54.5%	8.9%	
$R^2$ Fearful	2	58.4%	38.5%	19.9%	14.73***
France					
$R^2$ Overall	22	51.8%	56.7%	-4.9%	-2.01*
$R^2$ Bullish	19	27.3%	40.7%	-13.4%	-4***
$R^2$ Bearish	18	49.0%	43.9%	5.0%	1.10
$R^2$ Unfearful	21	38.6%	54.5%	-16.0%	-4.25***
$R^2$ Fearful	20	58.5%	57.1%	1.5%	0.37
Germany					
$R^2$ Overall	30	39.3%	47.7%	-8.4%	-2.35**
$R^2$ Bullish	16	29.2%	37.9%	-8.7%	-1.49
$R^2$ Bearish	16	51.0%	44.4%	6.6%	0.97
$R^2$ Unfearful	23	45.9%	50.8%	-4.9%	-1.22
$R^2$ Fearful	18	44.4%	45.5%	-1.0%	-0.16
Italy					
$R^2$ Overall	7	49.5%	53.4%	-4.0%	-0.88
$R^2$ Bullish	3	30.8%	39.1%	-8.3%	-2.96*
$R^2$ Bearish	5	39.6%	49.6%	-9.9%	-2.79**
$R^2$ Unfearful	6	40.9%	55.2%	-14.3%	-2.74**
$R^2$ Fearful	6	59.0%	60.7%	-1.6%	-0.25
Netherlands					
$R^2$ Overall	2	40.7%	64.4%	-23.7%	-0.73
$R^2$ Bullish	1	35.5%	59.1%	-23.5%	
$R^2$ Bearish	2	39.1%	44.2%	-5.1%	-0.24
$R^2$ Unfearful	1	66.4%	67.8%	-1.4%	
$R^2$ Fearful	2	46.6%	51.2%	-4.6%	-0.15
Spain					
$R^2$ Overall	6	65.6%	56.9%	8.7%	1.42
$R^2$ Bullish	4	58.0%	57.7%	0.4%	0.05
$R^2$ Bearish	4	64.8%	63.5%	1.3%	0.15
$R^2$ Unfearful	4	78.2%	70.4%	7.7%	0.58
$R^2$ Fearful	6	67.1%	61.9%	5.2%	3.38**
Switzerland					
$R^2$ Overall	3	54.8%	62.9%	-8.1%	-1.75
$R^2$ Bullish	3	18.0%	42.2%	-24.2%	-7.04***
$R^2$ Bearish	3	53.7%	39.2%	14.4%	0.79
$R^2$ Unfearful	3	35.6%	56.1%	-20.4%	-3.44**
$R^2$ Fearful	3	63.0%	63.9%	-0.9%	-0.2
United Kingdom					
$R^2$ Overall	17	32.8%	34.7%	-1.9%	-0.59
$R^2$ Bullish	12	21.8%	25.8%	-4.1%	-0.7
$R^2$ Bearish	16	32.0%	25.2%	6.8%	1.77*
$R^2$ Unfearful	13	31.4%	35.3%	-4.0%	-0.63
$R^2$ Fearful	15	40.2%	42.9%	-2.7%	-0.56

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Notes:  $R^2$  values are estimated for each firm using ordinary least square (OLS) regression model with  $r_{it} = \alpha_i + \beta_1 mkt_t + \epsilon_{it}$ . Thereby, N is number of regressions/firms, High is average of all regression  $R^2$  estimated based on firm weeks with high abnormal attention, Low is average of all regression  $R^2$  based on firm weeks with low abnormal attention, Delta is average difference in  $R^2$  between High and Low weeks' regressions and t-test is statistical significance of Delta. Bullish is based on firm weeks with positive market returns, Bearish is based on firm weeks with negative market returns, Unfearful is based on firm weeks with negative or stable changes in the level of the corresponding market volatility index, fearful is based on firm weeks with positive changes in the level of the corresponding market volatility index.

Table 18: Return predictability (per European country)

Dependent	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$r_t$	$r_t$	$r_t$	$r_t$	$r_t$	$r_t$	$r_t$	$r_t$
Sample	Belgium	France	Germany	Italy	Netherlands	Spain	Switzerland	UK
$r_{t-1}$	-0.148 (0.0250)	-0.0907*** (0.0182)	-0.0685*** (0.0210)	-0.0990* (0.0453)	-0.148 (0.0325)	0.0927 (0.113)	-0.146*** (0.00692)	-0.0283 (0.0282)
$ASV_{t-1}$	-0.0143** (0.000478)	-0.00653* (0.00332)	-0.00612 (0.00380)	-0.00359 (0.00747)	0.0291 (0.0131)	0.000602 (0.00559)	-0.0161 (0.00974)	-0.0211*** (0.00456)
$r_{t-1} \times ASV_{t-1} \times Bullish_{t-1}$	0.163* (0.0147)	0.164* (0.0860)	0.0523 (0.0630)	0.295 (0.310)	0.279 (0.774)	-0.351* (0.158)	0.450** (0.0850)	0.0851 (0.0851)
$r_{t-1} \times ASV_{t-1} \times Bearish_{t-1}$	0.302** (0.00668)	-0.0162 (0.0856)	-0.125 (0.0796)	-0.0157 (0.202)	0.423 (0.228)	-0.392** (0.113)	-0.600*** (0.0401)	-0.388*** (0.0630)
Constant	-0.587** (0.0232)	-0.0896 (0.0639)	-0.0493 (0.0324)	-0.303 (0.164)	-0.422 (0.0958)	-0.373** (0.0932)	-0.560* (0.145)	0.0746 (0.0655)
Controls	yes	yes	yes	yes	yes	yes	yes	yes
Observations	324	6,210	4,951	1,519	476	1,328	650	5,129
Number of firms	2	23	32	9	2	6	3	17
(Adjusted) R <sup>2</sup>	0.126	0.023	0.024	0.042	0.079	0.059	0.131	0.067

Firm-clustered standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Notes: Firm fixed-effect regression with clustered standard errors is applied.  $r_t$  is raw return in week t,  $r_{t-1}$  is lagged return by one week,  $ASV_{t-1}$  is logarithm of abnormal Google Search volume for firm name and "stock" in local language divided by median of last 8 week's Google Search Volume lagged by one week,  $r_{t-1} \times ASV_{t-1} \times Bullish_{t-1}$  is the interaction term of raw return and Google search volume lagged by one week if market return lagged by one week is greater or equal to zero and zero otherwise,  $r_{t-1} \times ASV_{t-1} \times Bearish_{t-1}$  is the interaction term of raw return and Google search volume lagged by one week if market return lagged by one week is smaller than zero and zero otherwise. Control variables include year dummies,  $Volat_t$  (historic volatility),  $Size_t$  (log market cap),  $BTM_t$  (book-to-market),  $InstOwn_t$  (institutional ownership),  $BASpread_t$  (bid-ask spread).