Can FEARS predict market returns?

VICTOR LÖTHNER

MARKUS PALSANDER

Bachelor Thesis

Stockholm School of Economics

2022

Can FEARS predict market returns?

Abstract

[This paper examines whether internet queries provided by Google Trends can proxy investor sentiment and thereby predict future market returns. By creating our own Financial and Economic Attitudes Revealed by Searches (FEARS) index using Internet Search Volume (SVI), we examine the period 2013-2022. We find FEARS (i) to not predict short term reversals, and (ii) to have a contemporaneous negative correlation with returns for multiple asset classes. These results differs from the study that we replicate and other earlier research within the field. We also find inconsistencies in the data which we believe to be the reason for why our period show different results compared to earlier periods. This paper contributes to the research of investor sentiment and more particularly the use of Google Trends as a proxy for investor sentiment.]

Keywords

[Investor sentiment, FEARS, Google Trends, Market returns, Limits to arbitrage]

Authors

Victor Löthner (24617)
Markus Palsander (24615)

Tutor

Christian Thomann, Researcher, Department of Finance

Examiner

Adrien d'Avernas, Assistant Professor, Department of Finance

Acknowledgments

We would like to thank the following people for helping us with this paper. Our tutor Christian Thomann for great insights and feedback about the process and methodology. Secondly, we would like to thank Sebastian Löthner for his support with the Python programming. All errors are on behalf of our mistakes.

Bachelor Thesis
Bachelor Program in Business and Economics
Stockholm School of Economics

1. Introduction

Investor sentiment, an area within the field of behavioral finance, tries to capture the overall sentiment of investors. There have been many studies that examine the impact of investor sentiment on financial markets. For instance, Baker and Wurgler (2006) find that sentiment broadly defined has significant cross-sectional effects on different types of stocks. Their measurement of sentiment is built using six different proxies that include NYSE share turnover and average first-day return on IPOs. While these proxies are valid measurements of sentiment, they work as an indirect proxy. Instead, internet queries have been proposed to be a more direct way of measuring household sentiment related to financial markets. One way of measuring household sentiment using internet queries has been by measuring their Google searches as they give more direct information about households' state of mind.

Our study is an extension of the Da, Engelberg, and Gao (2015) study and primarily contributes by examining whether the findings by Da et al. (2015) can be validated for an extended time period since a method that consistently can measure investor sentiment over time would be preferred. Such findings would also contribute to the robustness of Google Trends' ability to measure investor sentiment. By creating our own Financial and Economic Attitudes Revealed by Search (FEARS) index, using Internet Search Volume (SVI) which is provided by Google Trends, we use the FEARS index as a proxy for investor sentiment.

The study by Da et al. (2015) is conducted for eight years, ranging from January 2004 to December 2011. Our study includes nine years, ranging from January 2013 until December 2021. We have chosen an independent time period so their findings do not interfere with ours. Furthermore, we choose a slightly longer time period to check for more significance and include over 2000 observations. The study also captures the whole time period of the COVID-19 pandemic up until the study began.

The Da et al. (2015) study resulted in three major findings. They find that their own-created FEARS index (i) predicts short-term reversals, (ii) predicts temporary increases in volatility, and, (iii) predicts mutual fund flows out of equity funds and into bond funds. Moreover, they conclude these findings to, in general, be consistent with other research on investor sentiment.

Other studies have also managed to find that internet search queries measured by Google Trends have been able to predict market returns. In an earlier study by Da, Engelberg, and Gao, (2011) they were able to predict higher stock prices when SVI increased for different companies, measured by their stock ticker names. The same method with stock ticker names was also used by Wintoki, and Zhang (2011) to predict abnormal returns. By using their method, a trading strategy was formed by Bijl, Kringhaug, Molnár, and Sandvik (2016) which also yielded an abnormal return, but not after trading costs.
To measure household sentiment using internet queries, we use Google Searches since Google is by far the most used search engine on the internet with over 85%

of the market share in December 2021 according to Statista (2022). Furthermore, Google provides the service Google Trends which gathers data on the frequency of searches for different terms in different regions. The service Google Trends (https://www.google.com/trends) provides the SVI Index for every term which ranges from 0 to 100 depending on the number of searches on that term worldwide or from specific regions. The scale of 0 to 100 is relative to other dates within the given time period instead of giving an exact number of searches. As argued by Vozlyublennaia (2014) SVI reflects sentiment for retail investors rather than institutional investors might have more effective ways to gather information. While this is more true when the sentiment is measured using stock ticker names we anticipate that the majority of searches carried out reflect households rather than professional investors. In Figure 1 we show how the data for the two terms "PRICE OF GOLD" and "RECESSION" are displayed by Google Trends.

Figure 1
Interest over time for the term "PRICE OF GOLD"

Interest over time for the term "RECESSION"

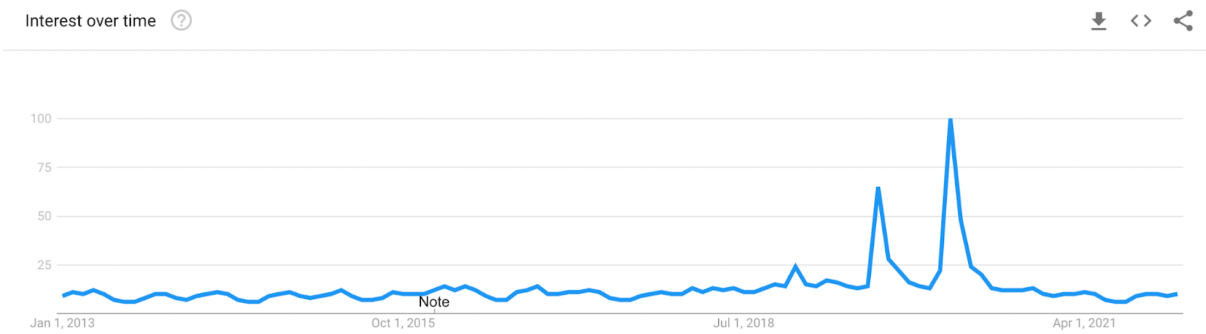


Figure 1 shows the monthly SVI for the terms "PRICE OF GOLD" and "RECESSION" as they are displayed by the service Google Trends (https://www.google.com/trends). The period plotted is 2013 January to 2021 December.

Da et al. (2015) stress the importance of identifying the relevant words that reveal the true investors' sentiment. By combining commonly known financial dictionaries, such as the Harvard IV-4 Dictionary and the Lasswell Value Dictionary, we generate a primitive list of 149 search words. Well known dictionaries are used in order to stay as objective as possible when choosing our terms (Tetlock, 2007). From the primitive word list, we access 1490 search terms. After identifying economic-related words that either reflect a positive or negative sentiment, as well as removing duplicates, we end up with a list of 155 terms including terms such as "GOLD", "RECESSION", and "ECONOMICS".

Moreover, we eliminate all the terms that are not economic-related as well as eliminating the terms with too few valid SVIs. The list then includes 137 search terms, which we then calculate daily log differences for. Due to comparable purposes, we also winsorize, remove seasonality, and remove heteroscedasticity through standardizing each term (Baker and Wurgler, 2006; Da et al., 2015). Lastly, we compute backward rolling regressions in order to assess which terms are of the highest interest.

Aside from our own-created FEARS index we also use the Economic Policy Uncertainty Index (EPU) developed by Baker, Bloom, and Davis (2013) to measure economic-political uncertainty, the commonly known market volatility index, VIX, from Chicago Board Operations Exchange (CBOE) that measures option-implied volatility on the S&P 500, and, also the Aruoba-Diebold-Scotti Business Conditions Index (ADS) developed by Aruoba, Diebold, and Scotti (2009) to measure business conditions during different times. All these indices are used as independent variables in our study.

We then use our FEARS index when testing on different asset classes. Our findings suggest that FEARS cannot predict short-term reversals, however, we find that an increase in FEARS correlates with a contemporaneous decline in market returns. For our time period, this appears to be true for multiple asset classes. Moreover, for bond index returns we find neither any reversals nor contemporaneous effects.

After relating our FEARS index to the asset classes, we do a final test where we consider limits to arbitrage. We create a high-low beta portfolio by calculating the spread from decile portfolios from data provided by CRSP. Moreover, we do a similar test for volatility, where we also create a high-low portfolio. Consistent with our previous findings, we do not find any reversals, although, we find contemporaneous correlations.

We compute multiple robustness checks to be certain of our results. We decided to use 30 search terms in our FEARS index throughout our previous calculations, but we also test with 25 and 35 words in our robustness checks. Additionally, we also test without any winsorization at all, as well as testing on different time periods. The results from our robustness checks show no sign of reversals. Da et al. (2015) also examine fund flows, which we do not. They conclude that FEARS can predict mutual fund flows from equity funds into bond funds. We have decided to restrict the scope of our study to market returns and asset classes.

2. Literature review

Behavioral finance contradicts the Efficient Market Hypothesis (EMH) and suggests a psychological view of financial models in order to explain market anomalies (Shiller, 2003). The research on investor sentiment emerged from the behavioral finance field, whereas Barberis, Shleifer, & Vishny (1998) define investor sentiment as the way that investors form their beliefs. Baker and Wurgler

(2007) suggest a similar definition of market sentiment as "the belief about future cash flows and investment risks that is not justified by the facts at hand". Baker and Wurgler (2007) determine that investor sentiment does affect stock prices, but the question is rather how one can measure investor sentiment and what its effects are. Moreover, they conclude that the stocks that are most affected by sentiment, are stocks that are difficult to arbitrage or value.

Investor sentiment can be measured in a variety of different ways. However, there are two methods that mainly have been used in order to measure investor sentiment. One method that empiricists have used, is market-based measures such as trading volume, closed-end fund discount, initial public offering (IPO) first-day returns, IPO volume, option-implied volatilities (VIX), and, mutual fund flows. The second approach used by some studies is a survey-based approach, where for example, one uses the Michigan Consumer Sentiment Index (MCSI), investment letters, UBS/GALLUP Index for Investor Optimism, or the Baker and Wurgler Index to measure investor sentiment (Brown and Cliff, 2005; Lemmon and Portniaguina, 2006; and Qiu and Welch, 2006; Baker and Wurgler, 2006, 2007). However, the second method mentioned suffers from the issue of respondents not needing to be truthful or accurate in their beliefs or opinions (Singer, 2002).

In addition to these two main methods used in order to proxy investor sentiment, more recent studies suggest a more unexamined method by studying micro-blogging sentiment or social network sentiment, which measures investor sentiment through internet usage by using Twitter, Facebook, Yahoo! Finance, or Google searches (e.g. Sprenger, Tumasjan, Sandner & Welpe, 2014; Zhang, Li, Shen & Teglio, 2016; Siganos, Vagenas-Nanos & Verwijmeren, 2017; Kim & Kim, 2014; Da et al., 2011; Da et al., 2015). These different methods of measuring sentiment show that there is no standardized way for how investor sentiment should be measured.

Investor sentiment does affect the stock market activity, where one example would be through asset valuation (Brown and Cliff, 2005). High investor sentiment can relate to that investors in the stock market being bullish about the current market outlook, and, is also connected to overconfidence (Liu, 2015; Odean, 1998). Furthermore, Barberis et al. (1998) suggest that due to the overconfident behavior of investors, momentum effects exist.

Investor sentiment is closely related to momentum strategies since momentum strategies build upon the foundation of buying stocks that rises. Jagadeesh and Titman (1993) examined the momentum effect and find evidence that momentum strategies outperform the S&P 500 index. Another study, also by Jagadeesh and Titman (2002), validated their findings and showed that over a three to 12-month period, the best-performing stocks continue to perform well the following three to 12 months. Moreover, they also find the opposite to be true, the stocks that perform worst continue to do so for the coming three to 12 months thereafter. Hence, their result provides an interesting reason for our study to try to validate the findings from Da et al. (2015) by measuring investor sentiment. Moreover, Yang & Zhou

(2015) find that investor sentiment can help explain market anomalies, for example, fire sales or limits to arbitrage. Research also suggests that limits to arbitrage can explain some anomalies that are fundamental-based, for example, post-earnings- announcement drift, accruals, and the cash holding effect (Li & Luo 2017; Mashruwala, Rajgopal, & Shevlin, 2006; Mendenhall, 2004).

In theory, e.g. De Long, Shleifer, Summers, Waldmann (1990) and, Shleifer and Vishny (1997) suggest that transaction cost and holding costs are reasons why arbitrageurs (rational investors) do not correct mispricings immediately when they occur. Thus, in the real world, arbitrageurs only exploit arbitrage mispricing when the profit exceeds the costs. Furthermore, the speed of corrections would be slower, implying that with high limits to arbitrage, the magnitude of mispricing would be larger.

Stambaugh, Yu, and Yuan (2012) suggest that firms with strong fundamentals are less likely to be underpriced since both optimistic, as well as pessimistic investors, are willing to invest in firms that have strong fundamentals during both high and low sentiment periods. Rather, it is a possibility that these firms instead are overvalued. Zhu, Sun, and, Yung, (2020) suggest that fundamentally strong firms that also have high arbitrage costs have lower future returns, in comparison with fundamentally strong firms that have high arbitrage costs when overpricing is corrected gradually. They also argue the opposite to be true, that fundamentally weak firms that have low arbitrage costs are less overpriced compared to fundamentally weak firms that have high arbitrage costs.

3. Data and methodology

3.1 Construction of FEARS index

Since our study is a replication of (Da et al. 2015) we try to keep as close as possible to their method when creating our FEARS index. Using Google Trends, SVI can be obtained for all terms that are searched by households. To find relevant search terms for our research we use the dictionary Harvard IV-4 and Laswell value Dictionary. By combining these two directories we can extract words based on categories from the Harvard IV-4 dictionary as well as filter those that display sentiment using Lasswell's value dictionary. The categories chosen are “Econ@” and “ECON” which means that the terms are related to economics and from them, we keep the ones that have a positive or negative sentiment tag. The method used to create the primitive word list makes sure we find terms using a method that limits hindsight bias which can arise when choosing the appropriate terms (Dzielinski, 2012). For instance, choosing words like “COVID” or “VACCINE” to better capture sentiment during the pandemic would suffer from such bias since investors at the time would not be able to anticipate covid to have the impact it had. Using this method we end up with a“primitive” list of 149 words. Using the primitive list we then find additional related words using Google Trends list of top ten related queries for each of those terms. These top 10 related queries can vary over time and we choose the top related ones for our time period 2013-2022.

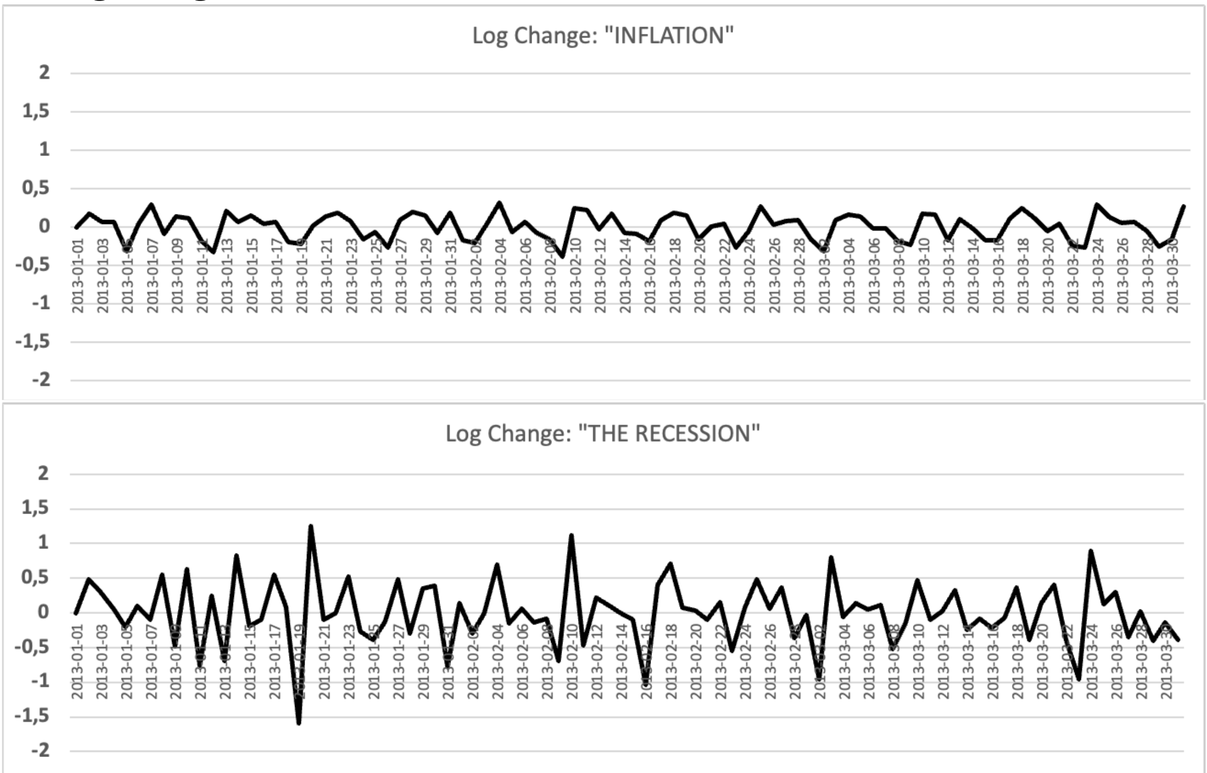
From the primitive word list, we get access to 1490 search terms that we have at our disposal. The next step is to find terms that are related to economics and finance to then create a final word list that will be used for further analysis and ultimately the construction of our FEARS index. To make the procedure more effective we make an initial manual sorting of the 1490 related search terms before downloading the data. In this sorting, we remove duplicates as well as terms that are clearly not related to finance or economics which leaves us with 155. An example of how we sort out terms that are not related to economics can be illustrated by some of the top 10 related search terms to the word "BANKRUPT" which yields the terms: "US BANKRUPT", "BANKRUPT COMPANIES", "BANKRUPTCY", "TRUMP BANKRUPT" "WHAT IS BANKRUPTCY" and "DETROIT". We choose to keep the first three terms as they are related to economics while the other terms are related to one individual; a question about the definition and a city, and are therefore removed. Finally, 18 terms with less than 1000 observations are removed which leaves us with 137 terms. These terms are then downloaded using a Python API for Google trends in increments of 3 months in order to receive daily data and then stitched together for the entire time period. Google Trends also gives us the option of downloading data from different regions like states, countries, and worldwide searches. Because our other independent and dependent variables are based on data from the United States alone we choose to only use data from the U.S.

Once we have downloaded the data for all the Search terms we want to calculate the change in interest over time. To do that we calculate the log change between day (t) and day (t-1) where (t) denotes the day our SVI is reported for and j denotes each search term.

$$\Delta SVI_{j,t} = \ln\left(SVI_{j,t}\right) - \ln\left(SVI_{j,t-1}\right)$$

When observing the changes in SVI for different terms we find that there are signs of seasonality in the data as illustrated by Figure 1 where we plot the log changes during the same quarter of 2013 for the two terms "INFLATION" and "THE RECESSION". For the former, we see a clear rise during Mondays followed by a progressive drop during the rest of the weeks. Comparing the log change for the two terms we also find that there are large differences in variance between them as well as some extreme values for the term "THE RECESSION". In order to address these issues, we decide to treat our data before finding the most relevant search terms to construct our FEARS index. First, we address the extreme values by winsorizing each **Δ**SVI by 5% (2,5% in each tail end). Secondly, seasonality is removed by regressing each **Δ**SVI on day of the week and month of the year dummy variables and keep the residual. Finally, we remove heteroscedasticity by standardizing each term. The standardization is done by scaling each term by its standard deviation to achieve a unit variance and standard deviation of one. What we are left with is an adjusted **Δ**SVI that is winsorized, seasonally adjusted, and standardized that we call **Δ**ASVI.

Figure 2
SVI log change for the terms "INFLATION" and "THE RECESSION"



These figures plot the log changes for the terms "INFLATION" and "THE RECESSION" during the same quarter (2013-01-01 to 2013-03-31). Both terms are also plotted on the same scale for comparability.

To then assess which search terms have the greatest correlation with market returns we do an expanding backward rolling regression for each **Δ**ASVI on market returns every six months. To find the most relevant terms for each 6-month period we, therefore, do a regression for all predating dates to find the most significant T-statistic. For example, to find the terms that will construct our FEARS index for the 6-month period of (1st of January 2019 - 30th of June 2019) we regress all terms on contemporaneous market returns for the time period (1st January 2013 - 31st December 2018). Because we need an initial window of 6 months to find the most relevant terms, our data period starts 1st of July 2013. We choose to use an expanding window for our regressions in order to find more significant results. From each regression, we rank all the terms (i=1) to (i=137) based on T-statistics from smallest to largest. In Table 1 we list the top 30 most negative significant terms that we find for a regression that covers our entire time period (1st of January 2013 - 31st of December 2021). What we find when looking at the entire time period is that the term by far with the lowest T-statistic is the term "GOLD PRICE" (T-statistic: -8,52), followed by "RECESSION" (T-statistic: -5,60) and "GOLD" (T-statistic: -5,26). Our findings are consistent with those of Da et al. (2015) in that we do not find any term that has a T-statistic greater than 2.5 that covers our entire time period while we find 8 terms that have a T-statistic lower than -2.5. Our findings are consistent with that of Tetlock (2007) in that negative words that show pessimism is more useful in measuring investor sentiment. Following that argument and because we find few positive terms with high T-statistic for our entire time period we only use terms with negative correlation with the market in our index. The FEARS index is constructed as follows:

$$FEARS_t = \sum_{i=1}^{30} R^i \,(\Delta ASVI_t)$$

Based on our results from each 6-month period we choose the 30 terms with the lowest T-statistics from each period to construct our FEARS index by simply calculating an equally weighted average of them. The choice of 30 terms is because it's the least amount of terms to eliminate idiosyncratic risk (Da et al. 2015). At the same time, we do not want to use too many terms as the significance becomes lower further down the list and since the terms have an equal impact on the index we want the more significant terms to not lose their impact. As a robustness check, we also calculate a FEARS index that contains 35 terms as well as one that contains 25 (see Section 7).

Table 1

	Search Term	T-Statistic
1	GOLD PRICE	-8,52
2	RECESSION	-5,60
3	GOLD	-5,23
4	FINANCIAL CRISIS	-3,03
5	US BANKRUPT	-3,03
6	BANKRUPT COMPANIES	-2,79
7	CRISIS	-2,74
8	INFLATION	-2,70
9	THE RECESSION	-2,42
10	2008 RECESSION	-2,39
11	TARIFF	-2,37
12	DEPRESSION	-2,07
13	DOLLAR	-2,06
14	THE DEPRESSION	-1,99
15	US RECESSION	-1,86
16	GREAT DEPRESSION	-1,85
17	SILVER	-1,83
18	CHARITABLE TRUST	-1,81
19	ECONOMY	-1,76
20	LAID OFF	-1,74
21	GOVERNMENT	-1,70
22	HOW MANY UNEMPLOYED	-1,69
23	BUSINESS EXPENSE	-1,66
24	SHORTAGE	-1,54
25	PROSPER	-1,50
26	BANKRUPT	-1,48
27	TARIFFS	-1,46
28	UNEMPLOYED RATE	-1,42
29	ECONOMICS	-1,41
30	ECONOMY RECESSION	-1,24

This table displays our top 30 terms out of 137 derived from Harvard-IV dictionary that has the most negative correlation with market returns for the time period 2013 January to 2021 December.

3.2 Other data

Most of the daily indices are collected from Capital IQ via Wharton Research Data Services (WRDS). We use the S&P 500 index as the main dependent variable due to its large sample of stocks and its weighting method. The S&P 500 (SPX) index is gathered from the CRSP via WRDS. Furthermore, we include the same exchange-traded funds (ETFs) as the Da et al. (2015) study for comparison purposes: the SPDR S&P 500 ETF Trust (ARCA: SPY), the Invesco QQQ Trust, Series 1 (NasdaqGM: QQQ), the iShares Trust - iShares Russell 1000 ETF (ARCA:

IWB), and the iShares Russell 2000 ETF (ARCA: IWM).[1] We use these ETFs in order to assure that our results are not driven by illiquid index component stocks. All four ETFs are gathered from CRSP via WRDS. We collect the return from the S&P U.S. Treasury Bond 10-Year Index gathered from Capital IQ via WRDS. The index comprise the most recent issued 10-year U.S. treasury notes or bonds. The index began on Septmeber 13th 2013 but provide values back to December 29th 1989. The S&P use backtesting to provide historical data before the start date of the index. However, the backtesting data regards a relatively short period of time in relation to our whole time period that is examined in our study. The treasury bond index differ from the one Da et al. (2015) use in their study, due to access limitations.

Moreover, we also use a number of independent variables, other than our own created FEARS index. VIX is used as an independent variable and estimates the volatility over the coming 30 days on the S&P 500 by calculating the implied volatility of options on the S&P 500 index. The VIX is commonly known as the "Fear Index" and an increasing value of the VIX implies more fear and uncertainty in the market. Hence, the index can be viewed as a proxy for investor sentiment. VIX was created by the CBOE, where we also collect the daily historical data for VIX. Furthermore, we also use the monthly MCSI in a separate test to examine whether there is correlation between the MCSI and Google Trends. The MCSI is collected from the Federal Reserve Economic Data (FRED) via the Federal Reserve Bank of St. Louis.

In order to measure real business conditions at high observation frequency, we use the ADS index, developed by Aruoba et al. (2009), as an independent variable in our tests. The ADS index is collected from FRED via the Federal Reserve Bank of Philadelphia, and measures a number of underlying economic indicators: weekly initial jobless claims; monthly payroll employment, monthly industrial production, monthly real personal income fewer transfer payments, monthly real manufacturing and trade sales; and quarterly real GDP. The underlying economic indicators are seasonally adjusted. The index is not cumulative and the average value for the index is 0, and the index can be used in order to compare business conditions during different times. Progressively more positive (negative) values indicate progressively better-than-average (worse-than-average) conditions.

As for our sample period, we see large spikes in the index during 2020 due to the COVID-19 outbreak, which is the largest value on the index during the whole period that the FRED database provides value for the index. We choose to include the values during the spikes year 2020 in our study to capture the extreme movements in market returns during this period.

---

[1] *The Invesco QQQ Trust, Series 1 (NasdaqGM: QQQ) was previously named PowerShares QQQ Trust (NASDAQ: QQQQ), which is the name used for the ETF in previous studies, such as Da et al. (2015)*

Baker et al. (2013) developed the economic policy uncertainty index (EPU) which is a news-based measure that measures policy-related economic uncertainty by three different components for the U.S. Firstly, the index for the U.S. is calculated by indexing ten large U.S. newspapers. Secondly, the index builds on reports from the Congressional Budget Office (CBO), by their list of temporary federal tax code provisions. Lastly, the third component of the index regards the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters. The data for the EPU index is collected from FRED via the Federal Reserve Bank of St. Louis.

Since the ADS index and the EPU index provides daily data (7-days), the weekends and holidays have not been included in our regression, since the daily equity market indices (5-days) do not provide any value for these days. Additionally, all the raw data has been collected on a daily basis, since the SVI index is updated daily (7-days). We have excluded values on weekends and holidays for the SVI index since values are missing for our dependent variables these days. Although, the MCSI index is an exception to our collection of daily values since the index only is measured on a monthly basis. Thus, the correlation with Google Trends is examined in a separate test.

We also use beta- and volatility-sorted portfolios in our calculations which are provided by CRSP via WRDS. The provided portfolios are divided into ten deciles where we calculate the return spread by calculating a high-low portfolio. The same procedure applies to both beta- and volatility-sorted portfolios. Additionally, we collect value-weighted and equally-weighted index returns from CRSP via WRDS, as additional asset classes.

3.3 Limitations

In our method, we have tried to stay as close as possible to the method proposed in the study we replicate. However, one limitation that we face in our work is the fact that we do not have access to the original dataset used by Da et al. (2015) which makes it difficult to know if our method is an exact replica of theirs (but for a different time period). While the method used to create the FEARS index and treat the data is simple to follow, one concern due to not having access to the original data file is related to the data gathering. For instance, Google Trends daily data can be downloaded for different time periods of up to 189 days before it switches to weekly data and for even longer periods only monthly data. When we compare data downloaded for 90 days and 189 days we find that for the same two terms, the average correlation for all SVIs we use is 0.828. The result means that the outcome can be very different depending on which time span one chooses to download the data. We ultimately go for the time period of 3 months when downloading the data since we suspect that is what has been done by Da et al. (2015) based on Figure 3 and 90 days sometimes being reported as the maximum span you can download daily data for (Kim, Lučivjanská, Molnár and Villa, 2019).

4. FEARS and Asset Returns

In the first section of FEARS and Asset Returns, we test for different asset classes as the dependent variable. In the second section, we do more tests that incorporate limits to arbitrage.

4.1 FEARS and average returns

The data that compose our FEARS index works as a proxy for investor sentiment which does not perfectly reflect the fundamental value of assets and the stock market. Different from informed investors, households often do not base their valuations based on future cash flows but instead other available information from news, or in our case, the internet. An effect of uninformed investors or so called noise traders is that stock returns unexplained by fundamentals are more likely to reverse back in the short run rather than those affected by fundamentals. One explanation for short term reversals is overreaction to information and cognitive biases i.e. investor sentiment (Da, Liu and Schaumburg, 2014). We therefore look for reversals due to changes in sentiment using the following regression:

$$return_{i,t+k} = \beta_0 + \beta_1 FEARS_t + \sum_m \gamma_m\, Control_{i,t}^m + u_{i,t+k}$$

Using our FEARS as an independent variable we begin our first test by conducting a regression for testing average returns. Other independent ($Control_{i,t}^m$) variables include the CBOE VIX index, daily change in ADS and daily changes in EPU. We also decide to include lagged asset class returns up to five lags. In the equation above (Return_{t+k}) denotes asset returns on time t+k. To spot potential reversals, we do a regression on contemporaneous returns then further test for the following days in order to see if any change remains or if it reverts back over the following five days from (k=1) until (k=5). Additionally, we test for cumulative returns over day one and two [Ret_{t+1}, Ret_{t+2}].

Our first test is done using the very broad equity market index S&P 500 in order to assess if household sentiment in the form of our FEARS index shows a contemporaneous correlation with market decline and reversal over the following days. Looking at the results from day (t) we see that an increase in our FEARS index has a negative contemporaneous correlation with market returns. This is not surprising as our FEARS index is constructed by the terms that have the most negative correlation with market returns. In Table 2 we show our result from the first regression. Recall that we standardized the SVIs that made up our FEARS index to have unit variance. However, since there is a lot of correlation between terms the standard deviation is not 1 for the final FEARS index but instead 0.248. This means that a change in one standard deviation for our FEARS index means a contemporaneous decline in the market of 10 basis points. The change is also significant at the 1% level. However, if we look at time (t+1) we only see a slight reversal which is not significant. Over the following days, we see the same pattern of insignificant predictability of the FEARS index. In Section 7 we will go over some of the possible explanations for why that might be. When looking at the EPU we see that it has predictive power when it comes to market returns for some of the following days returns. One explanation for the significance could be that it takes a while for households to adjust their expectations based on news-related policy uncertainty sentiment. However, the coefficients are very small which shows that it only predicts a minor change in market returns. As for the ADS, we know that there were some extreme changes during the COVID-19 pandemic which makes it better correlated with market returns.

Table 2

FEARS and S&P500 Index Returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Ret(*t*)	Ret(*t*+1)	Ret(*t*+2)	Ret[*t*+1,*t*+2]	Ret(*t*+3)	Ret(*t*+4)	Ret(*t*+5)
FEARS	–0.00397***	0.00038	–0.0000986	0.000283	0.0001517	0.0006745	–0.0004065
	(0.000894)	(0.009071)	(0.0000909)	(0.00118)	0.000909	(0.0009011)	(0.0009039)
VIX	–0.000233***	0.000053	0.000077**	0.000118***	(0.000074)**	0.00011***	0.00011***
	(0.0000317)	(0.0000324)	(0.0000324)	(0.0000422)	(0.0000324)	(0.000032)	(0.000032)
EPU	7.49e–07	–0.000012	0.0000111***	–5.92e–07	8.31e–07	8.36e–06 **	–8.98e–06**
	(3.73e–06)	(3.77e–06)	(3.70e–06)	(4.91e–06)	(3.77e–06)	3.74e–06	(3.75e–06)
ADS	0.00547**	0.00727***	0.00463***	0.01262***	0.00492***	0.00329*	0.002395
	(0.00177)	(0.00178)	(0.0018)	(0.00234)	(0.00170)	(0.00178)	(0.00179)
Ret(*t*)		–0.1638***	0.1116***	–0.0459	–0.08952	–0.07121***	0.04243*
		(0.0219)	(0.0219)	(0.0285)	(0.0219)	(0.0217)	(0.0218)
Ret(*t* - 1)	–0.1918***	0.0759***	0.00853	0.0839***	–0.0796***	0.0294	–0.0932***
	(0.0215)	(0.0221)	(0.221)	(0.0285)	(0.0221)	(0.0219)	(0.0220)
Ret(*t* - 2)	0.0664**	–0.0221	–0.0791	–0.0814***	0.00896	–0.0799***	0.1423***
	(0.0217)	(0.0219)	(0.0220)***	(0.0286)	(0.0220)	(0.0218)	(0.0219)
Ret(*t* - 3)	–0.255	–0.0850***	0.0113	–0.0720**	–0.0837***	0.14388***	–0.0903***
	(0.0218)	(0.0220)	(0.0220)***	(0.0287)	(0.0220)	(0.0218)	(0.0220)
Ret(*t* - 4)	–0.1165***	0.00141	–0.0836	–0.0770***	0.1471***	–0.0990***	0.1161***
	(0.0214)	(0.0229)	(0.0221)***	(0.0288)	(0.0221)	(0.0219)	(0.0220)
Ret(*t* - 5)	0.00154	–0.113***	0.165***	0.0492*	–0.1269***	0.1240***	–0.0320
	(0.0214)	(0.0216)	(0.0217)	(0.0282)	(0.0217)	(0.0215)	(0.0216)
Constant	0.00481	–0.000295	–0.000737	–0.000848	–0.000618	–0.00132	–0.00136
	(0.00059)	(0.000607)	(0.000601)	(0.000792)	(0.000608)	(0.000603)	(0.000605)
Observation	2,143	2,142	2,141	2,141	2,140	2,139	2,138
Adjusted R2	0.0870	0.0702	0.0663	0.0282	0.0673	0.0839	0.0788

This table shows the S&P 500 daily market returns as the dependent variable and regressed on the independent variables that are our FEARS index, control variables that include the VIX from CBOE, change in ADS, and change in EPU. Together with the control variables we also include lagged asset class returns on day (*t*+*k*) where (k=1) to (k=5), displayed on (1), (2), (3), (5), (6), and (7). We also include cumulative returns from [*t*+1,*t*+2]. *signifies significance at the 10% level, ** at 5%, and * at the 1% level.**

After testing for the S&P 500 we want to find out if we also find the same pattern of contemporaneous market decline and non-reversals on other asset classes. The rationale for testing other asset classes is to see if we achieve the same results as with the S&P 500 or if it is related to one index alone. Furthermore, we use highly liquid ETFs since reversals can be caused by liquidity shocks. Simply, when a fall in stock prices arises due to exogenous selling pressure on uninformed investors or public information that reduces investors valuation. Following that market makers require a higher expected return and prices rise during the following days (Campbell, Grossman, and Wang, 1993). Highly liquid ETFs will therefore isolate the effect of the change in sentiment reflected by our FEARS index. Other than equity portfolios we also test on a safer security that is reflecting bonds by using the S&P 500 ten-year maturity treasury bonds index.

In Table 3 we report the results from the different asset classes that we mentioned previously. In Panel A we report the results from the equally weighted and value-weighted CRSP index returns, in Panel B we show the results from the four different ETFs, and in Panel C we report the results from treasury bonds. Similar to our first tests we find that we do not find any significant reversal on the days following day *t*. We decide to only report the results for each asset class up until day (*t*+2) in accordance with Da et al. (2015) since we do not find any reversals for the first two following trading days and we do not anticipate any for the three days after that. In our result, we see the same pattern of non-significant reversals as in our first test. Because the ETFs and indexes we test are broad we see the same negative contemporaneous day (*t*) decline. A one standard deviation change corresponds with a contemporaneous decline of between 9.5 and 11.9 basis points for all indexes and ETFs which are significant at the 1% level. In most of our tests, we see a small reversal on day *t*+1 that is a fraction (about one-tenth) of the initial decline, however, it is not significant even at the 10% level which makes us unable to draw any conclusions from it. From the final test with treasury bonds, we do not find any signs of either significant reversals or significant contemporaneous change in bond returns.

Table 3
FEARS and returns to other asset classes

Panel A: FEARS and value-weighted and equally-weighted CRSP index returns								
	CRSP VW index returns				CRSP EW index returns			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]
FEARS	−0.00381***	0.000326	0.0000361	0.000354	−0.00424***	−0.000486	0.000469	−0.000042
	(0.000893)	(0.000905)	(0.000901)	(0.00121)	(0.00973)	(0.000975)	(0.000969)	(0.00137)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	2,138	2,137	2,136	2,136	2,138	2,137	2,136	2,136
Adjusted R2	0.0820	0.0660	0.0743	0.0304	0.0599	0.0647	0.0770	0.0502
Panel B: FEARS and ETF and RUSSEL 1000 and 2000 returns								
	SPY ETF returns				QQQ ETF returns			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]
FEARS	−0.00394***	0.000497	−0.000405	0.0000967	−0.00473***	0.0004301	−0.000685	−0.000225
	(0.000893)	(0.000904)	(0.000903)	(0.00119)	(0.00107)	(0.00108)	(0.00109)	(0.00143)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	2,138	2,137	2,136	2,136	2,138	2,137	2,136	2,136
Adjusted R2	0.0774	0.0626	0.0648	0.0261	0.0537	0.0413	0.0310	0.0168
	RUSSELL 1000 returns				RUSSELL 2000 returns			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]
FEARS	−0.00416***	0.000486	−0.000380	0.000114	−0.004814***	−0.000026	0.000086	0.00019
	(0.000902)	(0.000913)	(0.000913)	(0.00120)	(0.00114)	(0.00115)	(0.00115)	(0.00120)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	2,138	2,137	2,136	2,136	2,138	2,137	2,136	2,136
Adjusted R2	0.0845	0.0696	0.0686	0.0286	0.0586	0.0466	0.0460	0.0261
Panel C: FEARS and Treassury retunrs								
	Treasury returns							
	(1)	(2)	(3)	(4)				
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]				
FEARS	0.000314	−0.000136	0.000290	0.000157				
	(0.000355)	(0.000356)	(0.000351)	(0.000493)				
Controls	YES	YES	YES	YES				
Observations	2,123	2,122	2,121	2,121				
Adjusted R2	0.0045	−0.0004	0.0010	0.0013				

4.2 FEARS and limits to arbitrage

Before we come to the robustness of our results we want to do a final test in order to understand the predictability of our FEARS index. In order to do that we explore the concept of limit to arbitrage. The arbitrage pricing theory developed by Ross (1971) says that mispricings in the market create arbitrage opportunities that are then exploited and as a result remove the mispricing as the market adjusts. However, as put forward by Shleifer and Vishny (1997) there are limits to arbitrage that makes arbitrageurs more or less able to exploit arbitrage opportunities for some assets compared to others. For instance, assets with high volatility driven by sentiment investors, are considered unattractive if alpha does not increase proportionally to the increased volatility and in particular when fundamental risk is part of the volatility (Shleifer and Vishny, 1997). Furthermore, sentiment is costly to bet against and therefore harder to arbitrage any mispricings that resulted from it. Market Beta is also a feature that can have an impact on whether or not it can be hard to arbitrage away mispricings. High beta has this feature since an investment manager benchmark against the market would like to exploit mispricing in assets with betas close to 1. Also, high beta stocks are prone to more speculation and institutional constraints like bench-marking that make them unattractive to arbitrageurs (Baker, Bradley, and Wurgler, 2011). Because high beta stocks are less likely to be arbitraged away we anticipate a larger contemporaneous decrease as a result of an increase in sentiment.

To test this we create a high-low beta portfolio calculated from the spread between the highest and lowest beta portfolios that are formed in decile by CRSP using all stocks traded on NYSE and AMEX. We also test for the volatility by using the spread between the high and low volatility from the CRSP decile sorted, total volatility portfolios that use daily stock returns to calculate the total volatility.

In Table 3 we show the result from our regressions with beta and total volatility. Like previous tests, we see that when FEARS increase we see a contemporaneous decrease in returns on day (*t*) but we fail to find reversals on the following day. Worth noting is that there is a higher contemporaneous decrease in returns when FEARS increases on day (*t*) compared to previous tests. A change in one standard deviation in our FEARS index corresponds with a decrease of 12.3 basis points on day (*t*). Compared to S&P 500 with 9.5 basis points, there is a quite substantial difference from the one observed with a hi-lo beta portfolio. However, compared to that of Russell 2000 the difference is much smaller. When we look at the hi-lo total volatility portfolio we can see that we also find a contemporaneous decrease in return when FEARS increases on day (t). This also confirms our prediction that high total volatility stocks will be more affected by sentiment than stocks with low total volatility. On the other hand, we do not find any reversals for hi-lo total volatility on the days following a rise in FEARS.

Table 4
FEARS and limits to arbitrage

	Beta				Total Volatility			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret(*t*)	Ret(*t*+1)	Ret(*t*+2)	Ret[*t*+1,*t*+2]	Ret(*t*)	Ret(*t*+1)	Ret(*t*+2)	Ret[*t*+1,*t*+2]
FEARS	–0.004956***	–0.000895	0.000280	–0.00067	–0.003372**	–0.001546	0.001356	–0.00021
	(0.001574)	(0.001582)	(0.001587)	(0.02319)	(0.001341)	(0.001340)	(0.001348)	(0.002002)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	2,138	2,137	2,136	2,136	2,138	2,137	2,136	2,136
Adjusted R2	0.0166	0.0118	0.0062	0.0124	0.0224	0.0273	0.0162	0.0393

This table shows how our FEARS index correlates to daily returns of a hi-lo beta spread portfolio and a daily hi-lo total volatility spread portfolio over day (t), (t+1), (t+2), and cumulative returns during [t+1,t+2]. Controls include VIX from CBOE, change in EPU, and change in ADS and lagged returns up to five lags (k=1) to (k=5). *,, and *** denote coefficients are significant at 10%, 5% and 1% significance.**

4.3 Robustness checks

Because the results we find in our previous sections vastly differ from the ones found in Da et al. (2015) we carry out multiple robustness checks in order to solidify our findings. When we constructed our FEARS index we decided to go for 30 terms as our standard, which we then used in our previous regressions. As shown in Table 1 however, we can see that there is a fast decrease in significance among the ASVI displayed. Therefore, we want to test the robustness of our result by using two other FEARS indexes that are constructed in the same way as our original but with five fewer and five more terms that make up them ($FEARS^{35}$ and $FEARS^{25}$). Because we do not find reversals in our data we focus our robustness on the actual sentiment components in the form of our FEARS index since the other data with one exception up until now have been the same as in Da et al. (2015).

Another decision that impacted our FEARS index was the choice to winsorize each SVI by 5% (2.5% in each tail end) when adjusting to ASVI. Because winsorization reduces the extreme values, one concern could be that because of it, each of our ASVI could lose some of its predictive power. Our decision to winsorize is also made as a result of information that was not available if we were to create the index in real-time. This results in a forward-looking bias where we treat the data that are available to us ex-post. In order to account for this, we also construct a FEARS index where we refrain from winsorizing each SVI before continuing with any other adjustments like standardizing and removing seasonality. Next, we also check how our index performs in different subsamples in order to see if the predictability changes during different time periods.

For instance, during the early stages of the pandemic, we witnessed a sharp downfall in the market which was not followed by an immediate reversal during the following business days. There is a possibility that such a downfall in the market without immediate reversals could have an impact on our results since we fail to find any reversals the following day (t). In order to account for this, we test a time period that excludes the pandemic to see if our results change. We also decided to test for another time period that is even shorter since there is a potential that Google Trends could have lost predictive power gradually over time. Since we know that Da et al. (2015) find reversals during their time period and we, therefore, shorten our time period even further to be closer to theirs in order to eliminate possible loss of predictability later in our time period. A third period is tested that limits our time series to only the final three years. Recall that we in our construction of the FEARS index used a backward rolling regression to assess the most significant correlation with the market. Because the last period has the largest sample period, these terms are also most significant in terms of negative correlation. Finally, we decide to include a test that only includes our FEARS index as an independent variable to see if the index alone predicts reversals and if the control variables account for these.

In Table 5 we report the results from our robustness check. In Panel A we show the results from our two robustness FEARS indexes. We see that an increase in FEARS^{25} correlates with a contemporaneous decline in S&P 500 returns that is greater than that of our regular FEARS index. This is in line with our predictions since we have a cutoff point at 25 which includes more negative correlated terms, recall this from the construction of our FEARS index and Table 1. Similarly, an increase in FEARS^{35} correlates with a marginally lower contemporaneous decline in market returns.

In Panel B and C we show the results from the tests that involve the three subsections of our entire time series. We notice that when we limit the period to before 2020 and the pandemic the contemporaneous correlation loses some of its significance and is only significant at the 10%. When we limit the time period even more and check for only about 1000 observations from the beginning of 2013 to

July 2017 all contemporaneous correlation is lost. Therefore, we can estimate that our earlier period does not have more predictive power than later. However, some of the lost significant contemporaneous negative correlation can be due to fewer observations and the backward rolling regressions not including enough data. At the same time we do observe an increase in correlation with contemporaneous market decline on day (*t*) for our third-time series that only looks at the last third years. One interpretation could be that sentiment has a larger correlation with negative market returns during time periods with market instability like during the pandemic. On the other hand, our last time periods have the terms with the best fit thanks to the backward rolling regression which could increase the correlation during our time period.

In Panel 3 we show the result from our FEARS index that is constructed without winsorization. Here we find that an increase in FEARS has a larger negative correlation with market returns on day (*t*) compared to when we winsorize. In Appendix Table 6, we show the list of the 30 terms with the most negative correlation with market returns for these non winsorized ASVIs. We observe that the number one term "GOLD PRICE" has a more negative t-statistic than compared without winsorization see Table 1. Other than that change, the significance of each term remains relatively similar to that of our original ASVI.

Finally, in Panel D in Table 5 we show the results from our regression that only includes our FEARS index as an independent variable. We find that a larger contemporaneous market decline when FEARS increase which is due to the control variables in other tests explaining some of that decline. Throughout all our robustness checks including the one without control variables, we do not find any signs of significant reversals during the following days.

Table 5
Robustness check

Panel A: Construction	FEARS 25				FEARS 35			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]
FEARS	–0.004367***	0.000184	–0.000126	1.82e–06	–0.003957***	0.000495	0.000047	0.000532
	(0.000864)	(0.000891)	(0.000881)	(0.001156)	(0.000902)	(0.000915)	(0.000971)	(0.001194)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	2,143	2,142	2,141	2,141	2,143	2,142	2,141	2,141
Adjusted R2	0.0899	0.0445	0.0663	0.0141	0.0872	0.0703	0.0663	0.0283
Panel B: Subsection of timeseries	Before 2020				Before July 2017			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]
FEARS	–0.001354*	–0.000307	0.000990	0.000689	–0.000639	–0.000482	0.000356	–0.000053
	(0.000799)	(0.000834)	(0.000833)	(0.00116)	(0.000995)	(0.00103)	(0.00103)	(0.001434)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,637	1,637	1,637	1,637	1,008	1,008	1,008	1,008
Adjusted R2	0.0077	0.0030	0.0074	0.0107	0.0837	0.0005	0.0071	0.0146
Panel C: Subsection and No Winsorization	2019 to 2022				No Winsorization			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]
FEARS	–0.009328***	0.000260	–0.013675	–0.001117	–0.004376***	0.000344	–0.000062	0.000280
	(0.001855)	(0.001913)	(0.001927)	(0.0240)	(0.000907)	(0.000922)	(0.000924)	(0.001203)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	757	756	755	755	2,138	2,137	2,136	2,136
Adjusted R2	0.1770	0.1542	0.1397	0.0513	0.0890	0.0705	0.0662	0.0283
Panel D: FEARS without controls								
	(1)	(2)	(3)	(4)				
	Ret(t)	Ret(t+1)	Ret(t+2)	Ret[t+1,t+2]				
FEARS	–0.00430***	0.00105	–0.000521	0.000483				
	(0.000930)	(0.000935)	(0.000935)	(0.001193)				
Controls	NO	NO	NO	NO				
Observations	2,143	2,142	2,141	2,141				
Adjusted R2	0.0106	0.0107	–0.0003	–0.0004				

This table shows FEARS and different control tests that include two FEARS indexes that use the top 25 and 35 negative, most significant terms respectively. The table also shows our FEARS index and shorter time periods that include before 2020, before 2017, and 2019 to 2022. Finally, we display the results from FEARS without winsorization and without controls. All tests use S&P 500 as the dependent variable. *, and *** denotes significance at the 10%, 5% and 1% significance.**

5. Discussion

Based on our non-significant results, which differ from the Da et al. (2015) study, we here provide a discussion around the two main plausible reasons for what we believe may be the main reason(s). Firstly, we discuss the measurement and the methods of measuring investor sentiment. Secondly, the reliability and the predictive power of Google Trends are discussed.

5.1 Measuring investor sentiment

One explanation for why our results do not show validity or support short term reversals might be connected to the way of proxying investor sentiment. As mentioned previously, there are different approaches that include many different ways to proxy investor sentiment within those approaches. There is yet no standardized method for measuring investor sentiment, which explains the use of the many different methods that are being examined. We focus our study on using Google Trends to measure sentiment, where one proxy investor sentiment is based on different communication platforms and news on the internet. Other studies on microblogging sentiment also, for example, focus on proxy investor sentiment by examining Twitter or even music sentiment through Spotify as shown in a very recent study, where market reversals were found after an initial change in market returns as a result of a change in music sentiment. (Sprenger et al., 2014; Zhang, et al., 2016; Edmans, Fernandez-Perez, Garel, Indriawan, 2021).

Research shows that investor sentiment affects market returns (Baker and Wurgler, 2007). One theory that would explain our non-significant results for short-term reversals would be that this no longer is true. However, we find it unlikely that this is the case since this would imply that eminent research would be outdated and momentum strategies would no longer appear to be true. We agree with Baker and Wurgler's (2007) reflection regarding the question of how one can measure investor sentiment. Although, we want to extend this question further to whether investor sentiment even can be measured consistently over time, at all. It might be the case that different proxies work well during different time periods, but there is also a possibility that there is no proxy that will ever grasp the investor sentiment with significance.

While Da et al. (2015) highlights that there are concerns about the reverse causality in the prediction model and cannot conclude that sentiment causes return tomorrow. They find reverse causality unlikely thanks to their findings of positive reversals. We on the other hand, do not find any reversals and therefore reverse causality could be more of a risk. At the same time, we share their view in that it is unlikely that investors anticipating higher returns would search for negative terms.

5.2 Reliability of Google Trends

As we have discussed in the previous section there is a lot of evidence for investor sentiment affecting market returns and the different ways of measuring that impact. Our second explanation for why our tests do not show signs of predictability, therefore, relies on us using Google Trends as our proxy for the sentiment. There is no standardized way of using Google trends when doing predictions. In our study, we have shown one way that the data can be used but there are many more ways. For instance, one common practice when handling Google trends data is to use normalized data like in Joseph, Babajide Wintoki, and Zhang, (2011) which we do not include as it is not mentioned in the Da et al. (2015). Furthermore, methods have been proposed that improve the method for constructing a FEARS index using Google Trends. As argued by Yang, Dong, James, and Xu (2019) one way of improving the construction of a FEARS index using Google Trends is to weight the search terms based on impact rather than take an equally weighted average of all the 30 terms. This would make sense since as we saw in Tables 1 and 6 there is a big difference in how much an increase in a specific term correlates with a decline in the market. Furthermore, another concern with Google trends is that it can show some inconsistencies as we have already discussed a bit in the limitations part. As reported by Da et al. (2011) there is also a difference in data depending on when the data is downloaded where terms downloaded from different points in time had a 97% correlation. Taken together there are both measurement issues and different methods when it comes to using Google Trends to predict market movements and sentiment.

While the inconsistencies in the SVIs mentioned above may result in a difficulty to find the exact same result as another study, we believe the data overall should be accurate enough to give relevant information of the interest for a specific search term for any given time. When it comes to the optimal weighting of each SVI in our FEARS index we do recognize that there are improvements to be made, however, we do not believe this would lead to any finding of reversals. In our results, we find that the increase in FEARS correlating with a contemporaneous decline in returns is consistent over our tests with a significance of 1%. Furthermore, in our robustness test, we make adjustments that lead to a greater contemporaneous decline in market returns yet no increased sign of any reversals. Therefore, optimizing the construction of our FEARS index would lead to the finding of a greater contemporaneous decline yet we doubt that it would increase the chance of finding any reversals.

In our tests, we use controls that include VIX as well as ADS, EPU, and lagged market returns. One might therefore assume that these controls account for the reversals from our FEARS index. Because macroeconomic factors as well as previous returns can have a significant impact on investor sentiment (Vozlyublennaia, 2014). However, when not including any of the lags we are still not able to find any reversals which also confirms that our FEARS index does not predict reversals in the stock market. We, therefore, come to the conclusion that Google Trends is inconsistent as a proxy for sentiment rather than investor sentiment not having an impact or predicting market returns. Furthermore, when testing the correlation between the monthly MCSI and monthly SVI for the term "RECESSION" we find that the correlation coefficient is –0.1351 which means there is no correlation between the two and we show their relation in the Appendix (Figure 2). When comparing that to the time period used in Da et al. (2015) we find the correlation to be 0.862 which supports their finding which is almost identical (0.858).

Other studies find that using Google Trends as a measurement of sentiment has shown to be able to predict market returns (Da et al. 2011, 2015; Vozlyublennaia 2014; Joseph, Babajide Wintoki, and Zhang, 2011; Dzielinski 2012). However, with the exception of Da et al. (2015), these studies have used weekly data which in theory should have less ability to explain sentiment according to Vozlyublennaia (2014) since the destabilizing effect of attention from noise traders would be more noticeable on shorter time horizons. On the other hand, weekly data intuitively has the potential to be more accurate than daily data as it delivers a SVI based on more data and a longer period. Many previous studies also get their result from time periods ranging from 2004 to 2013 which means that our independent time period does not have to reflect theirs which is what we find. There is also some evidence that for later time periods Google Trends could not predict market returns as shown by Kim et al. (2019) who found that Google Trends could predict trading volume and volatility but not returns for the Norwegian stock market. Finally, almost all previous studies on our topic have used SVIs for stock ticker names to predict future market returns. With the support from the lack of correlation between SVI and MCSI, it is, therefore, more likely that terms we use and that are unrelated to specific company names have lost predictive power and we, therefore, do not find reversals or predictability in our data. One reason why this could be is that terms such as "RECESSION" and "THE DEPRESSION" now, more so than before, reflect interest and information gathering about such topics rather than investor sentiment. However, we do not have data that can confirm such explanations.

6. Conclusion

We use Google search terms like "RECESSION", "GOLD PRICES" and "FINANCIAL CRISIS" to create a FEARS index and investigate its prediction of market returns from 2013 to 2022. We find that an increase in FEARS correlates to a contemporaneous decline in market returns and no signs of reversals over the following days, both when we control and do not control for past asset class returns, CBOEs VIX, EPU, and ADS. This effect remains for different asset classes like broad ETFs, value- and equally-weighted index portfolios but not for bond index returns where neither reversals nor any contemporaneous effect is found. When testing for limits to arbitrage we find that for high-low spread sorted beta portfolios, an increase in FEARS correlates with a greater contemporaneous decline in returns compared to market returns (S&P 500) while not showing signs of reversals. We do not dispute the findings of Da et al. (2015) for their time period but rather question the predictability and reliability of Google Trends for later time periods. Contrary to their study, using our data we find that there is no correlation between the MCSI and the SVI "RECESSION". In light of this finding, we come to the conclusion that SVIs that are not related to stock tickers and company names have become worse at predicting market returns. However, the explanation why is not found. Our findings open for more research on Google Trends ability to reflect investor sentiment and predict market returns. A practical takeaway is that investors wanting to use strategies that involve sentiment from internet queries should do so with caution about the reliability of the data.

References

Aruoba, S., Diebold, F. and Scotti, C., 2009. Real-Time Measurement of Business Conditions. Journal of Business & Economic Statistics, 27(4), pp.417-427.

Baker, S. R., N. Bloom, and S. J. Davis. (2013). Measuring economic policy uncertainty. Working Paper, Stanford University.

Baker, M., Bradley, B., and Wurgler, J., 2011. Benchmarks as Limits to Arbitrage: Understanding the Low-Volatility Anomaly. Financial Analysts Journal, 67(1), pp.40-54.

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

Baker, M. and Wurgler, J., (2006). Investor Sentiment and the Cross-Section of Stock Returns. The Journal of Finance, 61(4), pp.1645-1680.

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

Bijl, L., Kringhaug, G., Molnár, P. and Sandvik, E., 2016. Google searches and stock returns. International Review of Financial Analysis, 45, pp.150-156.

Brown, G. and Cliff, M., 2005. Investor Sentiment and Asset Valuation. The Journal of Business, 78(2), pp.405-440.

Campbell, J., Grossman, S. and Wang, J., 1993. Trading Volume and Serial Correlation in Stock Returns. The Quarterly Journal of Economics, 108(4), pp.905-939.

Da, Z., Engelberg, J. and Gao, P., 2014. The Sum of All FEARS Investor Sentiment and Asset Prices. Review of Financial Studies, 28(1), pp.1-32.

Da, Z., Engelberg, J. and Gao, P., 2011. In Search of Attention. The Journal of Finance, 66(5), pp.1461-1499.

Da, Z., Liu, Q. and Schaumburg, E., 2014. A Closer Look at the Short-Term Return Reversal. Management Science, 60(3), pp.658-674.

De Long, J., Shleifer, A., Summers, L. and Waldmann, R., 1990. Noise Trader Risk in Financial Markets. Journal of Political Economy, 98(4), pp.703-738.

Dzielinski, M., 2012. Measuring economic uncertainty and its impact on the stock market. Finance Research Letters, 9(3), pp.167-175.

Edmans, A., Fernandez-Perez, A., Garel, A., and Indriawan, I., 2021. Music sentiment and stock returns around the world. Journal of Financial Economics, Forthcoming.

Jegadeesh, N. and Titman, S., 2001. Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. The Journal of Finance, 56(2), pp.699-720.

Jegadeesh, N. and Titman, S., 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. The Journal of Finance, 48(1), pp.65-91.

Joseph, K., Babajide Wintoki, M. and Zhang, Z., 2011. Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. International Journal of Forecasting, 27(4), pp.1116-1127.

Kim, S., & Kim, D. (2014). Investor sentiment from internet message postings and the predictability of stock returns. Journal of Economic Behavior & Organization, 107, 708–729.

Kim, N., Lučivjanská, K., Molnár, P. and Villa, R., 2019. Google searches and stock market activity: Evidence from Norway. Finance Research Letters, 28, pp.208-220.

Lam, F. and Wei, K., 2011. Limits-to-arbitrage, investment frictions, and the asset growth anomaly. Journal of Financial Economics, 102(1), pp.127-149.

Lemmon, M. and Portniaguina, E., 2006. Consumer Confidence and Asset Prices: Some Empirical Evidence. Review of Financial Studies, 19(4), pp.1499-1529.

Li, X. and Luo, D., 2016. Investor Sentiment, Limited Arbitrage, and the Cash Holding Effect. Review of Finance, p.2141–2168.

Linyi Yang, Ruihai Dong, Tin Lok James Ng, and Yang Xu. 2019. Leveraging BERT to Improve the FEARS Index for Stock Forecasting. In Proceedings of the First Workshop on Financial Technology and Natural Language Processing, pages 54–60.

Liu, S. (2015). Investor sentiment and stock market liquidity. Journal of Behavioral Finance, 16(1), 51–67.

Mashruwala, C., Rajgopal, S. and Shevlin, T., 2006. Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs. Journal of Accounting and Economics, 42(1-2), pp.3-33.

Mendenhall, R., 2004. Arbitrage Risk and Post-Earnings-Announcement Drift. The Journal of Business, 77(4), pp.875-894.

Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. Journal of Finance, 53(6), 1887–1934.

Qiu, L., and I. Welch. (2006). Investor sentiment measures. Working Paper, Brown University.

Shiller, R., 2003. From Efficient Markets Theory to Behavioral Finance. Journal of Economic Perspectives, 17(1), pp.83-104.

Shleifer, A. and Vishny, R., 1997. The Limits of Arbitrage. The Journal of Finance, 52(1), pp.35-55.

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

Singer, E. (2002). The use of incentives to reduce nonresponse in household surveys. Working Paper, University of Michigan.

Sprenger, T., Tumasjan, A., Sandner, P., & Welpe, I. M. (2014). Tweets and trades: The information content of stock microblogs. European Financial Management, 20(5), 926–957.

Stambaugh, R., Yu, J. and Yuan, Y., 2012. The short of it: Investor sentiment and anomalies. Journal of Financial Economics, 104(2), pp.288-302.

StatCounter. (March 1, 2022). Worldwide desktop market share of leading search engines from January 2010 to January 2022 [Graph]. In Statista. Retrieved May 10, 2022, from https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

Tetlock, P., 2007. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. The Journal of Finance, 62(3), pp.1139-1168.

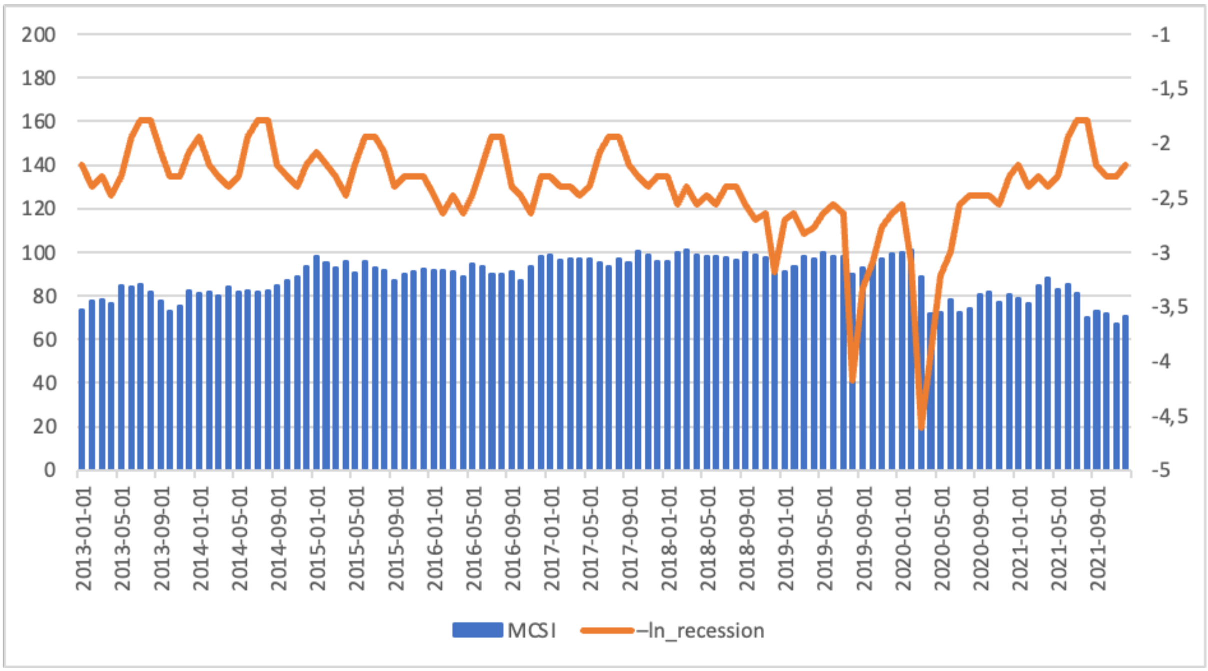
Vozlyublennaia, N., 2014. Investor attention, index performance, and return predictability. Journal of Banking & Finance, 41, pp.17-35.

Yang, C., & Zhou, L. (2015). Sentiment approach to underestimation and overestimation pricing model. Economic Modelling, 51, 280–288.

Zhang, W., Li, X., Shen, D., & Teglio, A. (2016). Daily happiness and stock returns: Some international evidence. Physica A-Statistical Mechanics and Its Applications, 460, 201–209.

Zhu, Z., Sun, L. and Yung, K., 2020. Fundamental strength strategy: The role of investor sentiment versus limits to arbitrage. International Review of Financial Analysis, 71, 101452.

Figure 2
"Recession" and MCSI



This figure shows the correlation between the monthly –(log) change for the SVI "RECESSION" and Michigan Consumer Sentiment Index (MCSI) for the time period 2013 January to 2021 December.

Table 6

	Search term	T-statistic
1	GOLD PRICE	-10,63
2	RECESSION	-6,97
3	GOLD	-4,11
4	US BANKRUPT	-3,23
5	FINANCIAL CRISIS	-3,06
6	INFLATION	-2,94
7	BANKRUPT COMPANIES	-2,68
8	DOLLAR	-2,58
9	CRISIS	-2,57
10	2008 RECESSION	-2,49
11	INTEREST	-2,46
12	THE RECESSION	-2,39
13	TARIFF	-2,27
14	ECONOMY	-2,19
15	DEPRESSION	-1,90
16	THE DEPRESSION	-1,86
17	US RECESSION	-1,80
18	GREAT DEPRESSION	-1,76
19	GOVERNMENT	-1,73
20	HOW MANY UNEMPLOYED	-1,72
21	TARIFFS	-1,70
22	SHORTAGE	-1,69
23	BUSINESS EXPENSE	-1,67
24	LAID OFF	-1,64
25	CHARITABLE TRUST	-1,54
26	UNEMPLOYED RATE	-1,47
27	ECONOMIC	-1,46
28	PROSPER	-1,35
29	ECONOMY RECESSION	-1,30
30	BANKRUPT	-1,29

This table displays our top 30 terms out of 137 derived from the Harvard-IV dictionary that have the most negative correlation with market returns when we do not winsorize the data. The time period is 2013 January to 2021 December.