THIS IS NOT FINANCIAL ADVICE

INVESTIGATING THE PREDICTIVE AND REACTIVE NATURE OF WALLSTREETBETS SENTIMENT

MELKER FALKENSTRÖM

THEO HASTINGS

Bachelor Thesis

Stockholm School of Economics

2025

This Is Not Financial Advice: Investigating The Predictive And Reactive Nature Of WallStreetBets Sentiment

Abstract:

In this paper, we develop a novel sentiment index (WSB) based on the frequency of bullish and bearish words occurring in finance-related comments on the WallStreetBets subreddit. We find that when WSB is high, same-day returns for the market, high minus low systematic risk portfolios and meme stock portfolios are low, and have been low for the three previous days. We find no evidence that WSB causes returns in the portfolios explored. Instead, WSB is found to be Granger caused by prior market and meme stock portfolio returns. When WSB is high, the VIX is likely to have increased in the previous days as well, suggesting that when markets expectations on future volatility is, and has been, high, sentiment is decreased on WallStreetBets. We further find a bilateral Granger causation relationship between WSB and VIX.

Keywords:

Big data, Sentiment, WallStreetBets

Authors:

Melker Falkenström (25816)
Theo Hastings (25727)

Tutors:

Ran Guo, Assistant Professor, Stockholm Business School

Examiner:

Ramin Baghai, Professor, Department of Finance

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

1 Introduction

In efficient market theories, securities are supposed to be priced based on all available information, reflecting the total future cash flows generated and their associated systematic risks. However, it has long been investigated whether other factors, such as investor sentiment, bounded rationality, or information asymmetries, that are not directly associated with the intrinsic value of the security can impact prices.

Since Baker and Wurgler (2006)[1] concluded that "when beginning-of-period proxies for sentiment are low, subsequent returns are relatively high for" small, young and high-risk stocks, many efforts have been concentrated towards developing new, innovative metrics to measure investor sentiment. Furthermore, since their work, the world has seen an increased adoption of digital technologies which has enabled the accelerated growth of software application usage. One such social network is Reddit, which has grown to 108+ million daily active users since its creation in 2005[1]. Reddit enables users to make posts and comments on various topics in different "subreddits" that tend to have a specific focus. For example the subreddit r/wallstreetbets, created in 2012, was established for stock and options interested individuals who were looking to place speculative bets and make short-term money. Today, WallStreetBets has grown into a vibrant community of 19 million subscribers[2], gaining popularity through a range of events in the recent decade, most noticeable in connection to the alleged coordinated short squeeze of investors shorting Gamestop and AMC Entertainment in 2021[3]. Even though WallStreetBets is widely known for its sarcastic tone and gravity towards memes and driving the popular term of meme stocks, the subreddit has become an influential forum that has caught the attention of retail and institutional investors alike. Investor sentiment proxies based on social media posts from communities such as WallStreetBets offer several advantages over both traditional measures (e.g., first-day IPO returns, implied volatilities, market surveys) and newer ones (e.g., Google search volume). Unlike many other sentiment indicators, WallStreetBets is uniquely focused on stocks and profit making, with a large and active user base generating real-time data. This allows for granular, high-frequency sentiment measurement. It also provides a rare combination of large volume and high concentration of stock-specific discussions. Users often express opinions on individual securities (e.g., "LVS will probably continue to rise and probably hits 70 at its peak"). By aggregating millions of such comments, it is possible to construct a sentiment index that reflects collective investor mood toward the market in fine-grained time periods.

Following our findings and by inspiration from other scholars, this paper focuses, not

[1] Reddit Inc., *The heart of the internet*, url: https://redditinc.com/, accessed May 11, 2025.

[2] https://www.reddit.com/r/wallstreetbets/, on WallStreetBets these are refered to as "degenerates"

[3] Jeffrey Neal Johnson, *WallStreetBets: How a Reddit Forum Shook Up Stock Market Dynamics*, url: https://www.nasdaq.com/articles/wallstreetbets-how-reddit-forum-shook-stock-market-dynamics, accessed March 9, 2025

only on sentiments impact on returns, but also volatility. Some scholars have furthermore questioned the assumption that sentiment is a leading indicator, instead suggesting that it may often reflect past market behavior. This aligns with our findings, and we therefore explore not only whether sentiment predicts returns and volatility, but also whether it reacts to them. The sentiment is proxied by mentions of bullish and bearish words on WallStreetBets and measures the daily bearish sentiment.

To begin with, we use bootstrapped regressions and granger causation analysis to explore the relationship between market returns and sentiment on WallStreetBets, we use the S&P 500 ETF as a proxy for the market. Building on the work by Stambaugh, Yu and Yuan (2012) [16], who found that stocks with high systematic risk performed worse following high sentiment realizations compared to those exposed to lower systematic risk, we investigate the relative WallStreetBets sentiment exposure of stocks with high versus low systematic risk. Therefore, a High Minus Low systematic risk (HML-risk) portfolio, based on the differences in returns from high beta stocks and low beta stocks is constructed. Furthermore, since discourse and trading activity among this community is often focused on so-called meme stocks, we investigate the nature of the relationship between frequently discussed meme stocks and sentiment on WallStreetBets. A dynamic portfolio, consisting of the 20 most mentioned stocks on WallStreetBets in the prior month, is developed to proxy for the performance of meme stocks. Finally, to investigate the relationship between volatility and the WSB index, an index is created using the volatility rank of the stocks included in the Meme portfolio as well as the HML-risk portfolio. The MEME VOL is calculated as the average of the volatility rankings in the dynamic meme stock portfolio, HML VOL is similarily calculated as the difference in average volatility rankings of the high and low beta stocks. We also include the VIX to represent a measure of future expected volatility.

To investigate the predictive power of the WSB index on SPY, the HML-risk portfolio, and the Meme portfolio returns as well as the HML VOL, Meme VOL and VIX, bootstrapped OLS regressions are conducted, controlling for past returns and additional control variables. To also capture the opposite direction of such predictive power, we include past as well as future returns for all regressions. Furthermore, bidirectional Granger causality tests are conducted between WSB and the six mentioned variables.

As for returns, we find no evidence that the WSB index can predict returns for any of the three portfolios, SPY, Meme and HML-risk. However the Meme portfolio as well as SPY granger causes WSB significantly and consistently for the first 18 lags. This suggests that past values of SPY and Meme portfolio returns can be used to predict WallStreetBets sentiment. As for volatility, we find that WallStreetBets precedes and Granger causes changes in relative volatility of meme stocks. Furthermore, we also find weak evidence for a reversal in this relationship, where the volatility rank of frequently discussed stocks first increases due to bearish sentiment and then decreases again one day later.

2 Literature review

The study of investor sentiment and its impact on financial markets has gained increasing attention over the last couple of decades. This trend has been driven by the growing availability of alternative data sources and advancements in sentiment analysis techniques.

Baker and Wurgler (2006) [1] found an inverse correlation between investor sentiment (INSE) and returns on speculative stocks, and drew the conclusion that INSE affects stock prices, however they concluded that further research was needed to both measure Sentiment and quantify the effects. Some proxies that have been used include market based measures such as IPO first day returns, Dividend Premiums and trading volume. Some literature has tried to come up with new proxies, for example the index developed by Cookson and Neissner (2020) [5] that used Stocktwits to measure investor sentiment and its predictive power on stock returns.

There has also been more newspaper focused research such as Tetlock (2007) [17], that found that the "Abreast of the Market" column in the Wall Street journal could have predictive power of daily stock returns between 1984 and 1999. Garcia (2013) [10], expanded this research and showed a similar result in the period 1905 to 2005. The proxy was based on the fraction of negative and positive words in the New York times financial news and found that, while controlling for well known time series patterns, the index had a stronger impact on predicting stock returns during recessions.

Research has also explored the connection between investor sentiment and market volatility, highlighting that sentiment not only impacts returns but may also drive short term volatility. One such contribution is the FEARS index developed by Da, Engelberg, and Gao (2015) [7], using Google search volumes for terms like "recession", "unemployment", and "bankruptcy" to proxy for household sentiment. Their findings suggest that increases in this index predict temporary reversals in returns, spikes in market volatility, and shifts in capital from equities to safer assets such as bond funds. These effects are especially short lived, aligning closely with theories of sentiment driven mispricing and noise trading. In particular, the results support the framework of De Long, Shleifer, Summers, and Waldmann (1990) [8], who argue that irrational noise traders can move prices away from fundamentals because arbitrage is limited when mispricing might worsen before correcting. In this view, sentiment can drive excess volatility and short term anomalies, challenging the assumption that markets quickly correct irrational behavior. Taken together, these findings emphasize the importance of including volatility, in addition to returns, when evaluating the market consequences of investor sentiment. It also introduces the need to explore how such a relationship behaves after a first reaction.

While much of the literature emphasizes the potential of sentiment to forecast returns and volatility, a number of studies instead suggest the reverse: that sentiment itself may largely be a byproduct of market conditions. Wang, Keswani, and Taylor (2006) [18], for instance, challenge the conventional assumption of sentiment as a leading indicator. They argue that previous research has often overstated sentiment's predictive power by omitting key lagged variables, particularly past returns and realized volatility. Using Granger causality tests on both daily and weekly sentiment proxies, including the ARMS index and various put–call ratios, they find that sentiment is frequently Granger-caused by returns and volatility, rather than the other way around. Once these lagged market variables are properly controlled for, the apparent forecasting power of sentiment diminishes significantly. This supports a reactive view of sentiment formation. However, some studies find evidence for the opposite direction of influence. Simon and Wiggins (2001) [15] demonstrate that market-based sentiment indicators, such as the VIX, put–call ratio, and trading index, can predict S&P 500 futures returns, often acting as contrarian signals where increased fear precedes positive returns. Similarly, Lee, Jiang, and Indro (2002) [13] find evidence that sentiment can influence both expected returns and volatility. Using a GARCH-in-mean method and the Investors' Intelligence index, they show that shifts in investor sentiment are contemporaneously associated with excess returns and that the magnitude of these shifts affects future volatility. According to Lee et al. (2002) [13], bullish sentiment tends to reduce conditional volatility and increase expected returns, supporting the idea that sentiment is a priced, systematic risk factor. These results further contribute to the view that sentiment can be predictive, rather than reactive. This divergence in empirical results highlights that the direction of the sentiment–market relationship remains unsettled and likely varies depending on the sentiment proxy used, the time horizon considered, and the characteristics of the asset class in question.

DeVault, Sias and Stark (2019) [9] expanded the research by investigating if it is individual or institutional investors that react to sentiments. Changes in sentiment should, according to DeVault et al. (2019)[9], in theory, induce a change in demand among sentiment investors. With a market clearing condition, in which every seller must be paired with a corresponding buyer, DeVault et al. (2019)[9] argue that they "can identify whether the traders captured by sentiment metrics are in aggregate individual or institutional investors, as changes in sentiment will be positively related to changes in demand from sentiment traders (i.e. demand shocks) for speculative stocks and inversely related to demand shocks from sentiment traders for safe stocks". In conclusion, they found that "an increase in sentiment is associated with both an increase in institutional investors' demand for risky stocks and, by definition, an associated decrease in individual investors' demand for speculative stocks." and DeVault et al. (2019)[9] drew the conclusion that sentiment metrics "capture institutional rather than individual investors' demand shocks." Furthermore, they stated that many of the sentiment metrics could be closely correlated with rational behaviour on the institutional side.

Building on this research we contribute by creating a novel sentiment index based on discourse on a highly speculative stock focused discussion forum, and exploring its relationship with the returns and fluctuations of the market, popular stocks on the forum and high minus low systematic risk portfolios.

3 Data description

3.1 WSB index

Due to lack of available, scrapable, social media sources we opted to use a pre-made dataset available on Academic Torrents that contains comments and submissions on the subreddit WallStreetBets since its creation in 2012 to June 2021. Academic Torrents is a network created for sharing of research data, and is an initiative that was launched by the Institute for Reproducible Research[4]. In total the dataset contains 49 290 700 comments and 1 900 000 submissions. Each comment and submission data point includes additional data apart from just the text that was posted, for example the name of the author, the timestamp of the post getting created, the timestamp of the author's account being created, if the author has a premium account etc. Each submission can receive multiple comments and the submission often tends to be relatively short (for example a short comment referring to a recent event) which explains the scale of obtained comments versus submissions. Due to these reasons, the analysis in this paper has exclusively been based on abstracted data extracted from the WallStreetBets comments.

A large share of the comments were not focused on securities and returns and were thus filtered away using 35 keywords as a basis for inclusion. The list of keywords includes terms such as buy, sell, calls, puts, bull, bear, SPY, and market, among others (see table 12 in appendix). These words were chosen to capture a wide range of investment related language used on retail investor forums like WallStreetBets, ensuring that the comments analyzed are indeed grounded in financial discourse. We also included several slang-based words to match the users of the platform. Since the data set is to be used for creating a sentiment index as well as serving as basis for a Meme portfolio, this method ensures that whenever, for example, Google is mentioned, it is indeed in the context of financial speculation. There is a sharp cut-off between having a restrictive list of keywords and still having a large enough dataset to be able to make a thorough analysis. A restrictive list would indeed give us a more reliable and accurate dataset, but would also lead to a much smaller dataset whereas a too extensive list of keywords would reduce the share of relevant comments. The size of the comment dataset and the large share of

[4]Institute for Reproducible Research, url: https://academictorrents.com/docs/about.html, Accessed March 2, 2025

comments centered on irrelevant topics resulted in us opting for a relatively restrictive list of keywords. Post-filtration the comment dataset consisted of 12 000 000 comments, thus even with this relatively restricted list of keywords, we managed to maintain approximately one fourth of the original dataset. See the full list of financial keywords words in table 12.

A list of bullish and bearish words was then constructed based on the words usage observed from sampled comments. This list of words contained 60 bullish words including bullish, pump, moon, FOMO, and breakout as well as 60 bearish words including bearish, crash, plummet, recession, and sell-off (see the full list in table 13 in appendix). The list was constructed based on the usage observed in a sample of comments and was selected for their frequent and contextually clear association with positive or negative market expectations.

In the selection of bullish and bearish words, a similar dilemma occurs as with the financial words list. A cut off between losing actual sentiment and catching false sentiment. We opted for a similar strategy as for the financial word list where a relatively restrictive list of bullish and bearish words were used.

A daily bearish index was then created by measuring the frequency of utilization of bearish and bullish words on a daily basis and setting the daily index as:

$$WSB_t = \frac{\#\, Bearish\, Words_t}{\#\, Bearish\, Words_t + \#\, Bullish\, Words_t}$$

In the years following the creation of WallStreetBets, user engagement remained relatively low with a generally low frequency of bullish and bearish words, to account for this the bearish index was analyzed in the period 2015-06-01 to 2021-06-30 (see figure 1). Through this method, we now have a bearish and bullish sentiment index for each day from 2015-06-01 and 2021-06-30.

The formation of this sentiment index includes several bias and noise problems. Firstly, the instances of parent comments are very apparent in the comment data from forums. The comment to which a reply responds is referred to as the parent comment, and this parent comment often plays a substantial role in determining the meaning of the reply, particularly in the context of determining sentiment. A response may appear to express a positive sentiment in isolation, but when interpreted in the context of a negative parent comment, it may instead reinforce that negative sentiment. We tried to minimize this bias by filtering the dataset by our lists of words. Our method ensures that each comment that is used to represent sentiment has to both include a finance-related word and have a clear sentiment. This removes all comments that are only confirmations and or refutations of parent comments such as “I agree” or “That is bullshit”. Furthermore, some comments may cause higher influence on other users through reputation and status. Since only a small minority of the comments had upvotes or downvotes we were unable to only include “supported” ideas. The noise and biases are many and it could be considered impossible to provide an exact sentiment index even when considering the quality of our data set. Looking into this problem with the help of Artificial Intelligence with tools like Large Language Models (LLM), deep learning etc could potentially reach closer. Early in our working process, LLMs were explored to potentially execute the sentiment analysis, but due to financial and time constraints, we opted to use other tools.

3.2 SPY

Daily prices and cash & cash equivalent distributions for the S&P500 ETF (SPY) were retrieved from the Wharton Research Data Services. Then, daily returns in the period 2015-06-01 to 2021-06-30 were calculated.

3.3 HML-risk portfolio

A list of all stocks in the Russell 1000 portfolio as of 2025 and corresponding returns was downloaded in the period 2010-01-04 to 2021-12-31. These returns as well as the Fama French Factors (Risk free rate and Excess market return) were retrieved from the Wharton Research Data Services. Excess returns for each stock was calculated for each recorded date as the return of the stock minus the risk free rate for that day. For each stock over the given period, the excess returns were regressed against the excess market return to calculate the beta value for each stock.

All stocks with available return data were then sorted based on their estimated beta values. The 30 stocks with the highest betas were grouped into a High-Risk portfolio and the 30 stocks with the lowest betas formed the Low-Risk portfolio. The average daily return was then calculated for each portfolio. Finally, the High minus Low Risk (HML-risk) portfolio was calculated by taking the difference in average returns between the High-Risk and Low-Risk portfolios.

Continuing, individual daily volatility ranks, indicating which decile for volatility individual securities belong to among 8300 securities on a daily basis, for the Russel 1000 stocks as of 2025 were collected through the SEC Market Information Data and Analytics System (MIDAS) dataset on WRDS for the period 2012-01-03 to 2021-06-30. Similarily to the HML-risk returns, the average daily volatility rank of the 30 highest (and lowest) stocks was computed. Finally the HML VOL portfolio was calculated as the difference in average volatility rankings of the high and low risk volatility rank portfolios.

3.4 Meme portfolio

The Meme portfolio is designed to represent a dynamic, WallStreetBets-inspired portfolio, composed of stocks that were most frequently discussed on the forum in the prior month. Given the trend-driven and volatile nature of discussions on the forum, it is essential that the portfolio adjusts dynamically to reflect current sentiment.

To include a wide range of potential stocks that could be discussed on WallStreetBets, stocks from the Russel 1000 portfolio as of 2025, combined with 42 meme stocks identified through WallStreetBets discourse, Quiver Quantitative[14] and Corporate Finance Institute[6], were used as a basis for the construction of the meme stock portfolio. Each stock was assigned a keyword that the authors deemed most likely to be used in association with the stock in WallStreetBets discourse. For instance, Wendys Arbys Group Inc is more commonly referred to as Wendys and not WEN which is its ticker. The 12 million filtered comments and the list of keywords for all securities were then used to identify the 20 most mentioned stocks on WallStreetBets in the month prior. Return data retrieved from WRDS was then used to calculate the equally weighted average return of the 20 most mentioned stocks in the previous month which forms what we in this paper will refer to as the Meme portfolio. Furthermore, individual daily volatility rank data collected through the SEC MIDAS dataset with WRDS was used to calculate the average daily volatility rank of the meme stock portfolio securities, forming the Meme VOL variable.

3.5 Control variables

To build on the methodology of Da et al. (2015)[7] we use the same control variables as in their paper. Namely the Daily Cboe Volatility Index (VIX), daily change in the Economic Policy Uncertainty Index (EPU), and daily change in the Aruoba-Diebold-Scotti (ADS) Business Conditions Index. This is made in an attempt to isolate the effect of WallStreetBets sentiment on meme stock performance, it's important to control for broader market and economic conditions that could influence both sentiment and returns. The VIX index is included to account for shifts in market volatility and investor risk appetite, which are especially relevant for speculative stocks. The EPU Index controls for fluctuations in policy-related uncertainty that may affect investor behavior regardless of sentiment. Lastly, the ADS Index captures changes in underlying business conditions at a high frequency. By including these controls, the analysis reduces the risk of omitted variable bias and helps ensure that any observed effect of WallStreetBets sentiment is not simply a reflection of wider macro-financial developments.

The Daily CBOE Volatility Index (VIX) is designed to measure expected 30 day volatility on US stocks, it is a benchmark index to measure the market's expectation of future volatility. The historical data was obtained from the Chicago Board Options Exchange official website. The index is calculated by CBOE using SPX options that expire on Fridays, including both standard monthly and weekly options. Only options with 23 to 37 days until expiration are included, and they are weighted to create a 30-day forward-looking measure of expected S&P 500 volatility. The index uses real-time bid/ask quotes, updated every 15 seconds, to compute ”spot” values throughout the trading day. Non-Friday expirations and SPX options outside the 23–37 day window are excluded from the calculation. See the summary statistics table for VIX in table 1.

The Economic Policy Uncertainty Index (EPU) is based on US Newspapers and measures the frequency of articles referencing words connected to Economy, Policy and Uncertainty. The historical data was obtained from the Federal Reserve Bank of Philadelphias official website. To measure policy-related economic uncertainty, the index is constructed based on three components. First, a newspaper-based index is constructed by quantifying the volume of articles related to economic policy uncertainty from 10 major U.S. newspapers, including the New York Times, Wall Street Journal, and Washington Post. This index captures media coverage of policy related uncertainty and is normalized over time. Second, data on scheduled expirations of federal tax code provisions, as reported by the Congressional Budget Office, is used to measure fiscal policy uncertainty. The expirations are dollar weighted and aggregated to reflect the anticipated uncertainty in the future tax code over a 10-year horizon. Third, forecast dispersion from the Federal Reserve Bank of Philadelphia’s Survey of Professional Forecasters is utilized to capture disagreement about future macroeconomic conditions. This includes uncertainty regarding the Consumer Price Index, federal expenditures, and state and local government spending. For our analysis we use daily change in the EPU index. See the summary statistics table of the EPU index in table 1.

The Aruoba-Diebold-Scotti (ADS) Business Conditions Index is a high-frequency indicator designed to track real-time developments in U.S. economic activity. The historical data was obtained from the Federal Reserve Bank of Philadelphias official website. It is constructed using six seasonally adjusted macroeconomic variables that vary in frequency: weekly initial unemployment claims, monthly payroll employment, industrial production, real personal income excluding transfers, real manufacturing and trade sales, and quarterly real GDP. These components are transformed into continuous, annualized growth rates (except for unemployment claims) and incorporated into a state-space model estimated via the Kalman filter. The ADS index is updated in real time, typically seven to eight times per month, based on data releases from official U.S. government agencies. The model assumes that ADS follows a daily first order autoregressive process and imposes a zero mean and unit variance. Additionally, the index is specified to move procyclically, with positive values indicating better than average business conditions. To account for pandemic-related data distortions, extreme COVID-19 values are excluded from parameter estimation but retained in the published index values. Similar to EPU, we use the daily change in the ADS index in our analysis. See table 1 for the summary statistics of the ADS index.

3.6 Remarks

Displayed in table 1 are summary statistics for HML-risk, Meme, SPY, WSB, ADS, VIX and EPU. VIX is used and displayed as the closing value of VIX. ADF stationarity results are displayed together with corresponding p-values. All seven variables are stationary with p-values below the 1% level. Finally, the summary statistics for the two volatility variables Meme VOL and HML VOL together with VIX is displayed in table 2. HML VOL is stationary with p-value $< 5\%$, but Meme VOL is not. However, changes in HML VOL and Meme VOL are both stationary with p-value below 1%.

Variable	Mean	Std	Min	25%	50%	75%	Max	ADF Stationary	ADF p-value
HML-risk	0.000	0.025	-0.185	-0.011	0.001	0.012	0.188	Yes	0.000
Meme	0.002	0.019	-0.135	-0.006	0.002	0.009	0.269	Yes	0.000
SPY	0.000	0.008	-0.051	-0.003	0.001	0.004	0.047	Yes	0.000
WSB	0.333	0.083	0.000	0.278	0.327	0.382	0.667	Yes	0.000
VIX	17.885	8.269	9.140	12.640	15.470	20.950	82.690	Yes	0.001
ADS	-0.318	3.659	-26.634	-0.377	-0.134	0.172	9.495	Yes	0.001
EPU	124.187	104.187	3.320	61.990	89.850	141.680	807.660	Yes	0.021

Table 1: Summary Statistics and ADF Test Results.
The table displays summary statistics for the returns of the three portfolios, Meme, SPY, and HML-risk, the bearish index WSB, as well as the three control variables, ADS, EPU, and VIX. All in the period between 2015-06-01 to 2021-06-30. The Augmented Dickey-Fuller (ADF) test is applied to test stationarity. In the table, mean, standard deviation (std), minimum value (min), first, second and third quartile, maximum value, as well as ADF results and ADF p-value is displayed.

Variable	Mean	Std	Min	25%	50%	75%	Max	ADF Stationary	ADF p-value
Meme VOL	3.681	0.984	1.900	2.950	5.500	3.500	4.300	No	0.157
HML VOL	3.285	0.735	1.049	2.779	3.263	3.767	5.691	Yes	0.047
Meme VOL Change	0.002	0.019	-0.135	-0.006	0.002	0.009	0.269	Yes	0.000
HML VOL Change	0.000	0.025	-0.185	-0.011	0.001	0.012	0.188	Yes	0.00
VIX	17.885	8.269	9.140	12.640	15.470	20.950	82.690	Yes	0.001

Table 2: Summary Statistics and ADF Test Results for volatilies.
This table reports summary statistics for Meme VOL, HML VOL, and VIX as well as the daily change in Meme VOL and HML VOL (Meme VOL Change, HML VOL Change). It includes the mean, standard deviation (Std), minimum, first quartile (25%), median (50%), third quartile (75%), and maximum values. The Augmented Dickey-Fuller (ADF) test for stationarity is also reported with the resulting p-values.

Figure 2 shows the standardized 30-day rolling averages of SPY returns and the WSB index over our time period (2015-06-01 to 2021-06-30). Intuitively, they tend to move

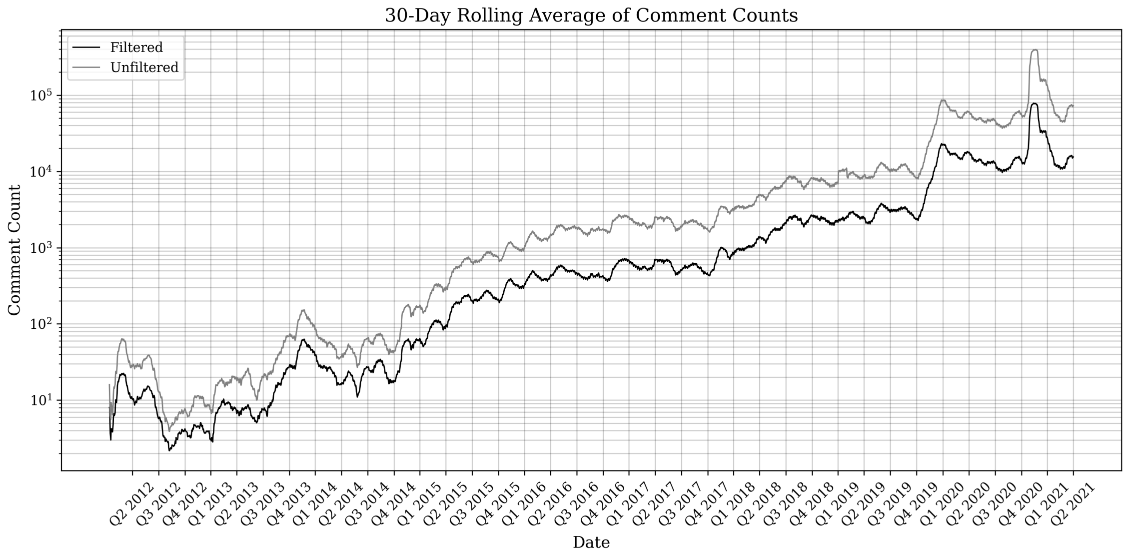

Figure 1: *The figure displays the 30 day rolling average comment count during the recorded time period in the dataset for the unfiltered comment dataset (which includes all collected comments) and the filtered comment dataset (only including comments mentioning finance and investment related words).*

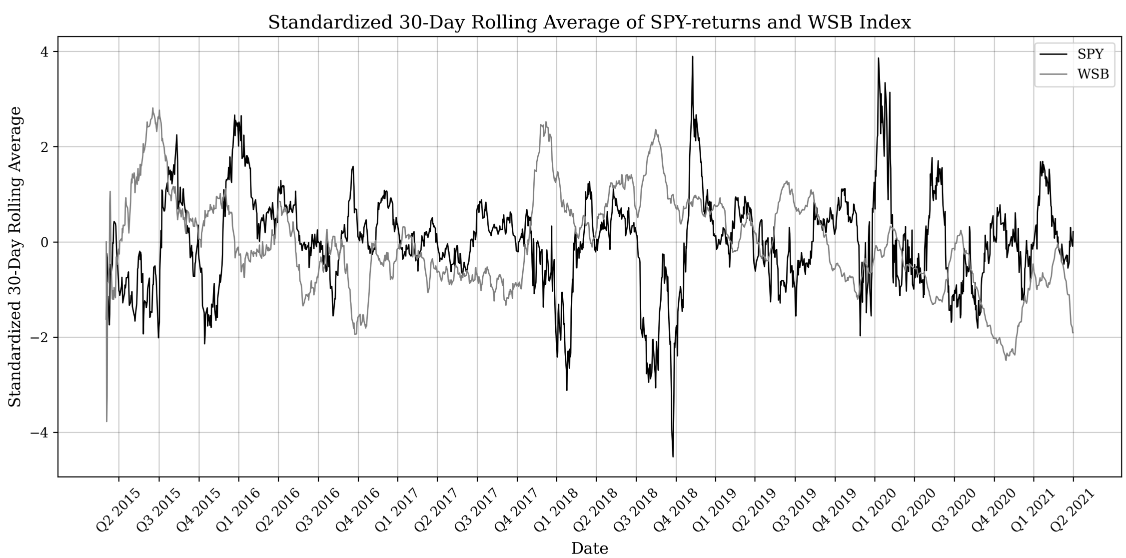

Figure 2: *The figure displays the standardized 30 day rolling average of SPY returns and the WSB index in the period 2015-06-01 to 2021-06-30.*

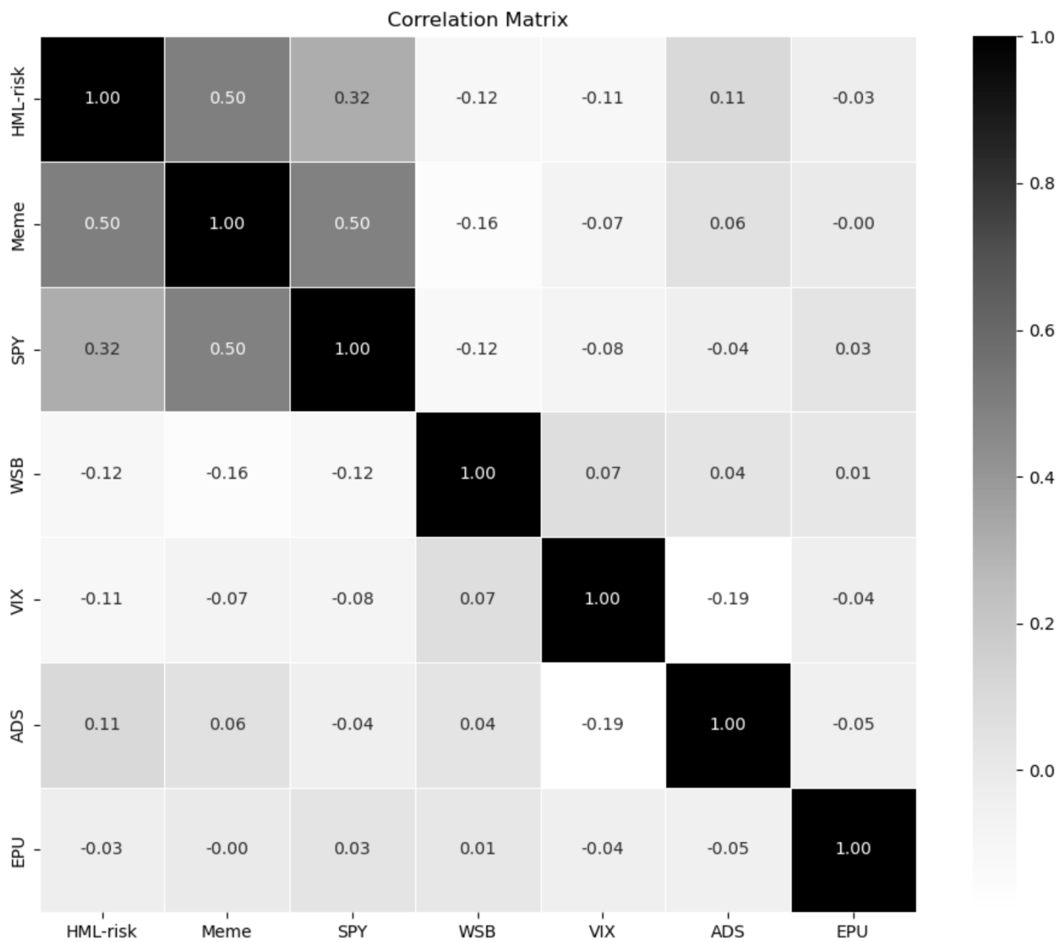

Figure 3: *The figure displays a correlation matrix for the three portfolios, Meme, SPY, and HML-risk, the bearish index WSB, as well as the three control variables, ADS, EPU, and VIX. Pearson correlation is applied.*

in opposite directions. Bearish WallStreetBets sentiment tend to coincide with periods of negative market returns, represented by SPY. These movements become especially evident in periods of high volatility, particularly in 2018 and early 2021.

The correlation matrix in figure 3 shows the relationships between the seven variables, HML-risk, Meme, SPY, WSB, VIX, ADS, and EPU. One can observe the strong positive correlation between the two portfolios created, HML-risk and Meme as well as SPY returns. Furthermore, the WSB index has a negative correlation with the three mentioned variables. VIX, ADS, and EPU show minimal correlations with the other variables, indicating little immediate interdependence.

4 Empirical Analysis

4.1 Returns analysis

In *The General Theory of Employment, Interest and Money* Keynes (1936) discussed the term ”Animal Spirits”. Koppl (1991) [12] reflects on his work and concludes that Keynes notioned that people were too ignorant to consistently behave in a rational manner, which could affect asset prices, arguing that peoples ignorance caused them to latch on to short term trading and speculation rather than long term investing, thus causing instability in the market.

If this is the case, it should be that an accurate measure of investors speculative view on the market could be correlated with, and perhaps even be used to predict future, asset price fluctuations.

To test whether and/or how proxies for sentiment on WallStreetBets are correlated with asset price fluctuations, we utilize bootstrapped ordinary least squares (OLS) regressions. The WSB index is used as the independent variable, and the returns of the S&P 500 ETF SPY are used as the dependent variable and serve as a proxy for market returns. Similarly, to Da et al. (2015)[7] we include changes in the ADS business conditions index and the economic policy uncertainty index as well as the CBOE VIX and lagged SPY returns as control variables. Based on the same logic applied by Da et al. (2015)[7] these control variables are included in order to isolate the effect of fluctuations in the WSB index by assigning business condition, volatility and economic policy effects to its direct causes as well as accounting for any auto regressive patterns in SPY returns. The regression formula applied thus becomes:

$$SPY_{t+i} = \alpha + \beta_1 \cdot WSBI_t + \beta_2 \cdot ADS_t + \beta_3 \cdot EPU_t + \beta_4 \cdot VIX_t + \sum_{j=1+max(-i,0)}^{5} \beta_{t+j} \cdot SPY_{t-j} + \epsilon_{ti}$$

Table 3 shows that the WSB index negatively predicts SPY returns for the three periods preceding as well as the contemporaneous period ($t-3$ to $t+0$), with the strongest effect at $t+0$ ($\beta = -0.001, p < 0.01$). In the periods post $t+0$, there are no significant results for the WSB index ($p > 0.1$ for $t+1$ through $t+5$). The results in table 3 thus indicate that the WSB index is negatively correlated with the same-day returns for SPY with a standard deviation increase in the WSB index corresponds to a decrease in the SPY returns of 0.0010. However, there are no significant signs of reversal trends in the following days SPY returns.

A lot of the discourse on WallStreetBets is not centered around general and broad-market indices; instead, a large share of the comments in our dataset concentrate on individual securities. Furthermore, WallStreetBets users, just like any other retail investor, face high search costs when trying to decide which common stocks to buy. According to Barber and Odean (2008)[2] such bounded rationality driven by high search costs causes retail investors to only buy stocks that have previously caught their attention. This suggests that users are more likely to consider buying stocks that have previously been mentioned in the discussions on WallStreetBets. Therefore, it could be that the securities that are discussed most frequently in a recent period on WallStreetBets, are more susceptible to be affected by sentiment. From a similar perspective, certain securities might exhibit characteristics that renders them more exposed to the sentiment of investors. For example, Stambaugh (2012) [16] found that stocks with high systematic risk performed worse following high sentiment measures compared to those exposed to lower systematic risk, as investors overestimated the performance of the high beta stocks that naturally should benefit the most from positive market factors. To further investigate the WSB index, similar regressions as those conducted for SPY were thus conducted using the developed Meme and High Minus Low Beta risk (HML-risk) portfolios.

The regression results in table 4 show a strong and statistically significant negative relationship between the WSB index and contemporaneous HML-risk portfolio returns ($\beta = -0.0027$ and $p < 0.01$). The same is true for lagged coefficients at $t-1$, $t-2$, and $t-3$ (e.g. $\beta = -0.0018$ and $p < 0.01$ at $t-2$, $\beta = -0.0015$ and $p < 0.05$ at $t-1$). However, as with the SPY portfolio, there are no significant results for subsequent lags (t + 1 through t + 5).

The regression results in table 5 highlight a strong and statistically significant negative relationship between the WSB index and same-day Meme portfolio returns, with a coefficient of $\beta = -0.0030$, significant at the 1% level. This suggests that a one standard deviation increase in WSB correlates with a 0.30 percentage point drop in Meme portfolio returns on the same day. Similarly, lagged coefficients at $t-1$, $t-2$, and $t-3$ are also negative and significant (e.g., $\beta = -0.002$ and $p < 0.01$ at $t-1$), indicating Meme portfolio returns having a predictive power on WallStreetBets sentiment. However, post t = 0, the relationship weakens, with coefficients at $t+1$ to $t+5$ being statistically insignificant

($p > 0.1$).

Similarly to the results in table 3, the results using the Meme portfolio returns and the High Minus Low Risk portfolio returns as the dependent variables indicates a significant contemporaneous relationship between the WSB index and same day returns for the respective portfolios. For example, it can be seen in table 5 that increases in the WSB index on a given day, is associated with lower returns for the 3 prior days as well as the same day. Similarly to the results in table 3 there are however no significant indications of a reversal effect in the following days. Similar results are also obtained in table 4 where the results indicate that an increase in the WSB index is associated with a lower return for stocks with high systematic risk, compared to their less risky counterparts.

Predictor	Ret(t-3)	Ret(t-2)	Ret(t-1)	Ret(t+0)	Ret(t+1)	Ret(t+2)	Ret(t+3)	Ret(t+4)	Ret(t+5)
const	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	0.0003	0.0002	0.0002
	(-0.0001, 0.0007)	(-0.0001, 0.0007)	(-0.0001, 0.0007)	(-0.0001, 0.0007)	(-0.0001, 0.0007)	(-0.0002, 0.0007)	(-0.0001, 0.0006)	(-0.0002, 0.0006)	(-0.0001, 0.0006)
WSB	-0.0005***	-0.0007***	-0.0008***	-0.0010***	0.0001	0.0002	0.0001	0.0001	-0.0000
	(-0.0009, -0.0002)	(-0.0011, -0.0003)	(-0.0012, -0.0004)	(-0.0015, -0.0006)	(-0.0004, 0.0005)	(-0.0002, 0.0007)	(-0.0003, 0.0005)	(-0.0003, 0.0005)	(-0.0004, 0.0004)
ADS	-0.0004	-0.0003	-0.0004	-0.0005	-0.0003	-0.0003	-0.0004	-0.0003	-0.0004
	(-0.0011, 0.0003)	(-0.0011, 0.0003)	(-0.0012, 0.0003)	(-0.0012, 0.0003)	(-0.0010, 0.0005)	(-0.0011, 0.0005)	(-0.0011, 0.0003)	(-0.0011, 0.0004)	(-0.0011, 0.0004)
EPU	-0.0002	0.0001	-0.0004*	0.0002	-0.0004*	0.0004*	-0.0004*	0.0006***	-0.0005**
	(-0.0007, 0.0004)	(-0.0005, 0.0006)	(-0.0009, 0.0000)	(-0.0004, 0.0007)	(-0.0009, 0.0000)	(-0.0000, 0.0009)	(-0.0008, 0.0000)	(0.0002, 0.0011)	(-0.0010, -0.0001)
VIX	-0.0005	-0.0005	-0.0005	-0.0008*	0.0004	0.0003	0.0002	0.0004	0.0003
	(-0.0013, 0.0003)	(-0.0013, 0.0003)	(-0.0015, 0.0003)	(-0.0016, 0.0000)	(-0.0005, 0.0012)	(-0.0006, 0.0011)	(-0.0006, 0.0009)	(-0.0004, 0.0012)	(-0.0004, 0.0010)
SPY lagged 1				-0.1108***	-0.0410	0.0216	-0.0643	0.0241	-0.0047
				(-0.1923, -0.0282)	(-0.1316, 0.0559)	(-0.0637, 0.1043)	(-0.1423, 0.0203)	(-0.0410, 0.0969)	(-0.0853, 0.0793)
SPY lagged 2			-0.1035***	-0.0699	0.0126	-0.0533	0.0121	0.0020	0.0678*
			(-0.1938, -0.0182)	(-0.1742, 0.0326)	(-0.0741, 0.0963)	(-0.1466, 0.0325)	(-0.0608, 0.0782)	(-0.0737, 0.0759)	(-0.0083, 0.1419)
SPY lagged 3		-0.1047**	-0.0628	-0.0161	-0.0536	0.0169	-0.0048	0.0715*	-0.0834**
		(-0.1911, -0.0225)	(-0.1638, 0.0385)	(-0.0999, 0.0699)	(-0.1397, 0.0298)	(-0.0610, 0.0906)	(-0.0810, 0.0703)	(-0.0072, 0.1527)	(-0.1621, -0.0120)
SPY lagged 4	-0.1004**	-0.0563	-0.0087	-0.0753	0.0106	0.0032	0.0689*	-0.0814**	0.0092
	(-0.1833, -0.0126)	(-0.1596, 0.0416)	(-0.0897, 0.0803)	(-0.1647, 0.0162)	(-0.0673, 0.0827)	(-0.0737, 0.0805)	(-0.0075, 0.1457)	(-0.1582, -0.0078)	(-0.0653, 0.0849)
SPY lagged 5	-0.0521	0.0018	-0.0638	-0.0072	-0.0080	0.0786*	-0.0866***	0.0154	-0.0066
	(-0.1534, 0.0475)	(-0.0819, 0.0767)	(-0.1573, 0.0312)	(-0.0807, 0.0667)	(-0.0812, 0.0713)	(-0.0030, 0.1580)	(-0.1574, -0.0184)	(-0.0584, 0.0876)	(-0.0825, 0.0646)

Table 3: *The table relates SPY daily returns to the WSB index. Building on a similar test done by Da , we use the contemporaneous SPY daily returns as well as the SPY daily returns over the next five days ($t+1$ to $t+5$) and the previous 3 days ($t-1$ to $t-3$). The independent variable is the WSB index. Furthermore, like Da , we use ADS, EPU, VIX and 5-day lagging SPY returns as control variables. We display bootstrapped 95% confidence intervals with 1000 iterations. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.*

Wang (2006)[18] found, as opposed to previous papers, that sentiment measures such as the OEX put-call trading volume ratio and the OEX put-call open interest ratio were caused by returns and volatility and not vice versa. Once controlling for lagged returns when using sentiment as a predictor for returns, the forecasting power of sentiment variables was extremely limited. To investigate the predictive power of sentiment for returns and realized volatility, Wang (2006)[18] ran Granger causality tests. Our results for the HML-risk portfolio and particularly the Meme portfolio and SPY suggests that sentiment, might, as Wang (2006) [18] found, be reactive to returns, and not vice versa. To further investigate the relationships between investor sentiment on WallStreetBets, the WSB in-

Predictor	Ret(t-3)	Ret(t-2)	Ret(t-1)	Ret(t+0)	Ret(t+1)	Ret(t+2)	Ret(t+3)	Ret(t+4)	Ret(t+5)
const	0.0004	0.0004	0.0004	0.0004	0.0005	0.0004	0.0004	0.0004	0.0004
	(-0.0008, 0.0017)	(-0.0009, 0.0016)	(-0.0009, 0.0016)	(-0.0009, 0.0017)	(-0.0008, 0.0017)	(-0.0008, 0.0017)	(-0.0009, 0.0017)	(-0.0009, 0.0018)	(-0.0008, 0.0017)
WSB	-0.0015***	-0.0018***	-0.0015**	-0.0027***	-0.0002	-0.0010	0.0001	-0.0001	-0.0000
	(-0.0027, -0.0005)	(-0.0030, -0.0006)	(-0.0028, -0.0002)	(-0.0039, -0.0015)	(-0.0015, 0.0010)	(-0.0022, 0.0003)	(-0.0011, 0.0013)	(-0.0013, 0.0011)	(-0.0011, 0.0012)
ADS	0.0028*	0.0030*	0.0025	0.0023	0.0027	0.0023	0.0019	0.0018	0.0017
	(-0.0003, 0.0058)	(-0.0003, 0.0061)	(-0.0010, 0.0060)	(-0.0010, 0.0058)	(-0.0007, 0.0062)	(-0.0012, 0.0058)	(-0.0015, 0.0055)	(-0.0014, 0.0051)	(-0.0016, 0.0050)
EPU	-0.0011	0.0005	-0.0013	-0.0006	-0.0008	0.0005	0.0007	0.0005	-0.0009
	(-0.0029, 0.0006)	(-0.0015, 0.0026)	(-0.0029, 0.0002)	(-0.0024, 0.0012)	(-0.0023, 0.0008)	(-0.0011, 0.0022)	(-0.0009, 0.0021)	(-0.0012, 0.0023)	(-0.0028, 0.0008)
VIX	-0.0018	-0.0016	-0.0013	-0.0020	0.0007	0.0007	0.0011	0.0016	0.0018
	(-0.0048, 0.0016)	(-0.0050, 0.0016)	(-0.0045, 0.0020)	(-0.0052, 0.0012)	(-0.0023, 0.0038)	(-0.0022, 0.0036)	(-0.0021, 0.0041)	(-0.0013, 0.0044)	(-0.0012, 0.0048)
HML-risk lagged 1				0.0631	0.0355	0.0119	0.0261	-0.0145	-0.1018**
				(-0.0436, 0.1652)	(-0.0643, 0.1338)	(-0.1049, 0.1335)	(-0.0663, 0.1183)	(-0.1078, 0.0868)	(-0.1972, -0.0065)
HML-risk lagged 2			0.0635	0.0135	0.0096	0.0119	-0.0168	-0.1001**	0.0703
			(-0.0428, 0.1692)	(-0.0838, 0.1079)	(-0.1168, 0.1396)	(-0.0898, 0.1089)	(-0.1112, 0.0828)	(-0.2009, -0.0053)	(-0.0220, 0.1664)
HML-risk lagged 3		0.0610	0.0156	-0.0080	0.0147	-0.0223	-0.1013**	0.0699	-0.0202
		(-0.0447, 0.1603)	(-0.0784, 0.1078)	(-0.1356, 0.1312)	(-0.0842, 0.1115)	(-0.1107, 0.0753)	(-0.2001, -0.0049)	(-0.0140, 0.1701)	(-0.1189, 0.0689)
HML-risk lagged 4	0.0573	0.0145	-0.0093	-0.0087	-0.0248	-0.1101**	0.0676	-0.0203	0.0139
	(-0.0481, 0.1501)	(-0.0883, 0.1075)	(-0.1381, 0.1140)	(-0.1124, 0.0971)	(-0.1210, 0.0734)	(-0.2080, -0.0130)	(-0.0265, 0.1722)	(-0.1158, 0.0715)	(-0.0719, 0.0962)
HML-risk lagged 5	0.0169	-0.0085	0.0035	-0.0400	-0.1020**	0.0598	-0.0173	0.0205	0.0330
	(-0.0880, 0.1076)	(-0.1378, 0.1158)	(-0.0929, 0.1081)	(-0.1321, 0.0592)	(-0.2061, -0.0093)	(-0.0293, 0.1531)	(-0.1112, 0.0758)	(-0.0709, 0.1124)	(-0.0542, 0.1108)

Table 4: *The table relates the HML-risk portfolio daily returns to the WSB index. The contemporaneous HML-risk portfolio daily returns as well as the HML-risk portfolio daily returns over the next five days ($t+1$ to $t+5$) and the previous 3 days ($t-1$ to $t-3$) are displayed. The independent variable is the WSB index and the HML-risk portfolio daily returns is the dependent variable. Furthermore, like Da , we use ADS, EPU, VIX and 5 day lagging HML-risk portfolio returns as our control variables. Bootstrapped 95% confidence intervals are displayed with 1000 iterations. *, **, and *** denote significance at the 10%, 5%, and 1% level respectively.*

dex, and the Meme portfolio, the HML-risk portfolio and SPY, six Granger causality tests are employed.

The Granger causality test is a standard econometric tool used to assess whether one time series contains predictive information about another. Granger causality evaluates whether past values of one variable improve the forecast of another variable, controlling for the target variable's own lagged values. One variable "Granger-causes" another variable if information about the past of the first variable helps predict the future of the second variable better than using only its own past data (Granger, 1969) [11]. Mathematically, the model can be expressed as:

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \beta_i Y_{t-i} + \sum_{i=1}^{p} \gamma_i X_{t-i} + \epsilon_t$$

Where Y_t is the dependent variable, X_t is the independent variable, p is the number of lags considered, α_0 is the intercept term, β_i are the coefficients for the lags of the dependent variable Y_t, γ_i are the coefficients for the lags of the independent variable X_t, and ϵ_t is the error term. The null hypothesis (H_0) for the Granger causality test is that X_t does not Granger cause Y_t, i.e., past values of X_t do not provide additional information to

Predictor	Ret(t-3)	Ret(t-2)	Ret(t-1)	Ret(t+0)	Ret(t+1)	Ret(t+2)	Ret(t+3)	Ret(t+4)	Ret(t+5)
const	0.0014***	0.0013***	0.0014***	0.0015***	0.0014***	0.0015***	0.0017***	0.0015***	0.0016***
	(0.0006, 0.0023)	(0.0004, 0.0021)	(0.0005, 0.0023)	(0.0006, 0.0024)	(0.0005, 0.0024)	(0.0006, 0.0025)	(0.0008, 0.0026)	(0.0006, 0.0024)	(0.0007, 0.0024)
WSB	-0.0018***	-0.0015***	-0.0020***	-0.0030***	0.0005	-0.0002	-0.0006	-0.0000	-0.0005
	(-0.0027, -0.0011)	(-0.0024, -0.0007)	(-0.0029, -0.0009)	(-0.0038, -0.0021)	(-0.0003, 0.0013)	(-0.0012, 0.0007)	(-0.0017, 0.0003)	(-0.0010, 0.0009)	(-0.0014, 0.0004)
ADS	0.0011	0.0011	0.0009	0.0009	0.0009	0.0008	0.0008	0.0006	0.0006
	(-0.0008, 0.0030)	(-0.0009, 0.0031)	(-0.0010, 0.0029)	(-0.0010, 0.0029)	(-0.0011, 0.0029)	(-0.0011, 0.0029)	(-0.0011, 0.0029)	(-0.0011, 0.0025)	(-0.0012, 0.0023)
EPU	-0.0005	0.0014*	-0.0018***	-0.0003	-0.0006	0.0010*	-0.0000	0.0012**	-0.0009
	(-0.0016, 0.0005)	(-0.0000, 0.0028)	(-0.0030, -0.0006)	(-0.0014, 0.0008)	(-0.0016, 0.0004)	(-0.0001, 0.0021)	(-0.0010, 0.0009)	(0.0001, 0.0023)	(-0.0022, 0.0005)
VIX	-0.0004	-0.0003	-0.0003	-0.0008	0.0013	0.0011	0.0013	0.0014*	0.0016*
	(-0.0024, 0.0018)	(-0.0024, 0.0019)	(-0.0023, 0.0016)	(-0.0030, 0.0014)	(-0.0005, 0.0030)	(-0.0009, 0.0030)	(-0.0004, 0.0030)	(-0.0001, 0.0029)	(-0.0001, 0.0031)
Meme lagged 1				-0.0541	0.1357**	0.0995*	-0.0611	0.0209	-0.1031*
				(-0.1764, 0.0943)	(0.0259, 0.2384)	(-0.0006, 0.2234)	(-0.1503, 0.0518)	(-0.0678, 0.1056)	(-0.1949, 0.0019)
Meme lagged 2			-0.0360	0.1244**	0.0969*	-0.0579	0.0041	-0.0996*	0.0496
			(-0.1589, 0.1321)	(0.0071, 0.2200)	(-0.0119, 0.2175)	(-0.1543, 0.0502)	(-0.0870, 0.0879)	(-0.1849, 0.0013)	(-0.0344, 0.1282)
Meme lagged 3		-0.0507	0.1257**	0.0723	-0.0639	-0.0086	-0.0794	0.0436	-0.0179
		(-0.1760, 0.1133)	(0.0177, 0.2404)	(-0.0351, 0.1799)	(-0.1561, 0.0355)	(-0.0811, 0.0714)	(-0.1669, 0.0163)	(-0.0394, 0.1211)	(-0.1029, 0.0654)
Meme lagged 4	-0.0448	0.1249**	0.0750	-0.0919*	-0.0128	-0.0913*	0.0562	-0.0203	0.0914*
	(-0.1740, 0.1173)	(0.0160, 0.2339)	(-0.0291, 0.1906)	(-0.1816, 0.0108)	(-0.0928, 0.0605)	(-0.1737, 0.0004)	(-0.0242, 0.1315)	(-0.1088, 0.0587)	(-0.0067, 0.2117)
Meme lagged 5	0.1217**	0.0816	-0.0785	-0.0185	-0.0841*	0.0675*	-0.0301	0.0975*	-0.0258
	(0.0111, 0.2270)	(-0.0223, 0.1916)	(-0.1716, 0.0198)	(-0.0915, 0.0588)	(-0.1674, 0.0079)	(-0.0033, 0.1438)	(-0.1074, 0.0568)	(-0.0035, 0.2075)	(-0.0993, 0.0560)

Table 5: *The table relates the Meme portfolio daily returns to the WSB index. The contemporaneous Meme portfolio daily returns as well as the Meme portfolio daily returns over the next five days ($t + 1$ to $t + 5$) and the previous 3 days ($t - 1$ to $t - 3$) are displayed. The independent variable is the WSB index and the dependent variable is the Meme portfolio returns. We control for ADS, EPU, VIX, and 5-day lagged Meme returns. Bootstrapped 95% confidence intervals with 1000 iterations are reported. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.*

predict Y_t. The alternative hypothesis (H_1) is that past values of X_t help predict Y_t. The test is performed by calculating the F-statistic for each lag. If the p-value is below the significance threshold (0.05), we reject the null hypothesis and conclude that X_t Granger causes Y_t.

The first test tests the hypothesis that the WSB index Granger-causes Meme portfolio returns and the second tests the hypothesis that the Meme portfolio returns Granger-causes the WSB index. This bidirectional approach is essential, as the relationship between retail investor sentiment and meme stocks may be complex and potentially reciprocal. The same is done for the HML-risk portfolio and SPY on their relationship with the WSB index.

The results in table 6 reveal a clear asymmetry in predictive power. At all tested lags (1 through 5), WSB fails to Granger-cause SPY returns, with p-values consistently above significance thresholds for all measures and F-values below one for all lags. In contrast, SPY significantly Granger-causes WSB across all lags, with p-values consistently below 0.05 for all lags and F-values above 3 for all lags. In addition, the significant results remains until lag 18 (see appendix table 13 and 14). These results suggest that movements in the broader market precede changes in WSB, but not the other way around.

Lag	Test	SPY → WSB	WSB → SPY	Meme → WSB	WSB → Meme	HML → WSB	WSB → HML
Lag 1	SSR F	4.1270**	0.0686	6.0778**	0.2482	0.8472	0.1224
	Chi²	4.1351**	0.0687	6.0897**	0.2487	0.8488	0.1226
	LR	4.1295**	0.0687	6.0776**	0.2486	0.8486	0.1226
	Param F	4.1270**	0.0686	6.0778**	0.2482	0.8472	0.1224
Lag 2	SSR F	4.5643**	0.4478	5.6617***	0.7206	1.7777	0.7388
	Chi²	9.1586**	0.8985	11.3606***	1.4458	3.5671	1.4824
	LR	9.1313**	0.8983	11.3186***	1.4451	3.5630	1.4817
	Param F	4.5643**	0.4478	5.6617***	0.7206	1.7777	0.7388
Lag 3	SSR F	4.2474***	0.4458	5.1119***	1.0866	1.6178	0.6606
	Chi²	12.8007***	1.3436	15.4063***	3.2748	4.8757	1.9909
	LR	12.7475***	1.3430	15.3293***	3.2713	4.8680	1.9896
	Param F	4.2474***	0.4458	5.1119***	1.0866	1.6178	0.6606
Lag 4	SSR F	3.7421***	0.2085	4.0171***	0.8130	2.0770	0.5531
	Chi²	15.0572***	0.8390	16.1635***	3.2711	8.3571	2.2256
	LR	14.9835***	0.8388	16.0787***	3.2676	8.3344	2.2240
	Param F	3.7421***	0.2085	4.0171***	0.8130	2.0770	0.5531
Lag 5	SSR F	3.3334***	0.1675	3.5422***	0.7023	2.1937*	0.4220
	Chi²	16.7879***	0.8434	17.8393***	3.5369	11.0478*	2.1252
	LR	16.6964***	0.8431	17.7360***	3.5328	11.0081*	2.1237
	Param F	3.3334***	0.1675	3.5422***	0.7023	2.1937*	0.4220

Table 6: *The table presents the results of Granger Causality tests between the WSB index and the three return based portfolios: SPY, Meme and HML-risk. For each pair of variables, we conduct bi-directional Granger causality tests over five different lag lengths (Lag 1 to Lag 5). each corresponds to one trading day. The reported statistics include the SSR-based F-test, the Chi-squared statistic, the Likelihood Ratio (LR) test, and the parameter-based F-test. Finally, * , ** and *** denotes significance at the 10%, 5% and 1% level.*

A similar result is true for the Meme portfolio. Through all five lags, the WSB fails to Granger-cause Meme portfolio returns. As for SPY, p-values are above 10% for all lags and all F-values are below 1. In contrast, the Meme portfolio returns do Granger-cause WSB across all lags, with p-values below 0.05 for all lags. We further observe higher F-values than those obtained with SPY in the same direction, especially at lag 1-3. In addition, the significance remains until lag 18 (see appendix table 13 and 14). These findings indicate that changes in meme stock activity precede shifts in bearish sentiment, but not the other way around (see table 6).

Finally, the results for the HML-risk portfolio shows, just as with the other two portfolios, insignificant results for the test if WSB granger causes HML-risk portfolio returns. However, it shows weaker significance levels for the opposite direction with only weak significant results for lag 5. For this test we also see the lowest F-values across the three portfolios (eg. SSR F = 0.8472 at lag 1 and SSR F = 2.1937 at lag 5).

Table 6 provides evidence that retail sentiment, as measured by the WSB index, is reactive to broader market conditions, SPY, and to meme stock movements, but does not seem to predict these variables. This indicates that speculative retail trading behavior is largely driven by shifts in the broader market and meme stock dynamics, rather than the other way around. Our findings from both the OLS regressions and the Granger causality tests supports that of Wang (2006)[18]. Once controlling for lagged returns when using sentiment as a predictor for returns, the forecasting power of sentiment variables is extremely limited. For meme stocks specifically, the results emphasize that retail sentiment, via WSB, may be caused by rather than causing meme stock returns. This is consistent with the results from the OLS regressions for the Meme portfolio (See table 5). These results, rather than aligning with the results of Da (2015)[7] is in line with the results of Brown and Cliff (2004)[4] who did not find evidence supporting the conventional idea that sentiment measures affect individual investors. Finally, similar to Brown and Cliff (2004)[4], we find that past returns seems to be a determining factor of the sentiment of today.

4.2 Volatility analysis

Using an overlapping generations model De Long (1990) [8] concluded that introducing irrational noise traders to the markets caused prices to fluctuate from their fundamental value to a higher degree, thus inflating volatility in markets that previously fluctuated to reflect available information. Sentiment changes among retail investors could, therefore, impact not only the prices, but also the volatility of securities. A long strand of empirical research supports this view. For example, using data from the 1990s, Brandt, Brav, Graham, and Kumar (2010) [3] found that stocks with low prices and low institutional ownership exhibited higher idiosyncratic volatility, thus supporting the view that idiosyncratic volatility can be inflated through increased retail investor trading activity. As the WSB index aggregates the sentiment of risk-on retail investors, we therefore aim to investigate the relationship between the WSB index and market volatility. Similarly to the returns section, as a lot of the discourse on WallStreetBets is centered on high risk and individual securities, we investigate the relationship between the WSB index and the volatility of the beta sorted high minus low risk and Meme portfolios.

To begin with, we use the CBOE volatility index (VIX) and OLS regressions to explore how the WSB index is related to market expectations on future volatility. As increases in the WSB index is designed to indicate an increased bearish sentiment among WallStreetBets investors we expect to find a positive relationship between WSB and closing VIX values. Similarly to the returns section we include lagged VIX values and changes in the ADS and EPU index to control for autocorrelation and changes in policy uncertainty or underlying business conditions.

Next, to investigate if the volatility of certain securities might be more or less exposed or correlated to the WSB index, we run OLS regressions where the WSB index, along with the control variables, are regressed against the daily changes in HML VOL and Meme VOL portfolios.

Predictor	Ret(t-3)	Ret(t-2)	Ret(t-1)	Ret(t+0)	Ret(t+1)	Ret(t+2)	Ret(t+3)	Ret(t+4)	Ret(t+5)
const	-0.0002	-0.0001	-0.0001	0.0002	0.0004	0.0008	0.0010	0.0015	0.0017
	(-0.0123, 0.0115)	(-0.0120, 0.0134)	(-0.0123, 0.0126)	(-0.0120, 0.0125)	(-0.0142, 0.0162)	(-0.0163, 0.0200)	(-0.0204, 0.0236)	(-0.0215, 0.0255)	(-0.0229, 0.0271)
WSB	0.0148**	0.0183***	0.0254***	0.0303***	0.0131	-0.0007	-0.0114	-0.0147	-0.0173
	(0.0017, 0.0295)	(0.0041, 0.0334)	(0.0103, 0.0429)	(0.0149, 0.0468)	(-0.0060, 0.0312)	(-0.0205, 0.0194)	(-0.0354, 0.0114)	(-0.0408, 0.0112)	(-0.0461, 0.0126)
ADS	-0.0315*	-0.0303	-0.0240	-0.0253	-0.0396**	-0.0565***	-0.0695**	-0.0800***	-0.0871***
	(-0.0700, 0.0044)	(-0.0682, 0.0078)	(-0.0621, 0.0088)	(-0.0609, 0.0070)	(-0.0768, -0.0050)	(-0.1051, -0.0121)	(-0.1304, -0.0158)	(-0.1421, -0.0240)	(-0.1537, -0.0261)
EPU	0.0003	-0.0057	0.0231***	-0.0093	-0.0010	-0.0131	-0.0139	-0.0160	-0.0131
	(-0.0143, 0.0155)	(-0.0215, 0.0106)	(0.0087, 0.0386)	(-0.0241, 0.0058)	(-0.0175, 0.0166)	(-0.0329, 0.0054)	(-0.0364, 0.0100)	(-0.0417, 0.0086)	(-0.0410, 0.0138)
VIX lagged 1				0.7930***	0.8431***	0.8704***	0.7676***	0.7855***	0.7248***
				(0.6536, 0.9336)	(0.7063, 0.9852)	(0.7165, 1.0365)	(0.6112, 0.9348)	(0.5898, 1.0240)	(0.5449, 0.9310)
VIX lagged 2			0.7894***	0.1962*	0.1788	0.0498	0.1440	0.0763	0.1306
			(0.6495, 0.9492)	(-0.0115, 0.4146)	(-0.0495, 0.4011)	(-0.1839, 0.2799)	(-0.1291, 0.4256)	(-0.2144, 0.3554)	(-0.1380, 0.4100)
VIX lagged 3		0.7841***	0.2028*	0.0218	-0.1114	-0.0193	-0.0616	-0.0165	-0.0953
		(0.6217, 0.9400)	(-0.0119, 0.4342)	(-0.1472, 0.2100)	(-0.2731, 0.0294)	(-0.2488, 0.2389)	(-0.3406, 0.2042)	(-0.2780, 0.2577)	(-0.3774, 0.1943)
VIX lagged 4	0.7791***	0.2002*	0.0300	-0.1202*	-0.0309	-0.0786	-0.0148	-0.1141	0.0447
	(0.6160, 0.9368)	(-0.0031, 0.4322)	(-0.1205, 0.1981)	(-0.2701, 0.0184)	(-0.1944, 0.1537)	(-0.3012, 0.1482)	(-0.2450, 0.2200)	(-0.3769, 0.1347)	(-0.2152, 0.2974)
VIX lagged 5	0.1903**	-0.0164	-0.0537	0.0773*	0.0653	0.0981	0.0595	0.1429	0.0459
	(0.0336, 0.3566)	(-0.1589, 0.1120)	(-0.1679, 0.0630)	(-0.0122, 0.1694)	(-0.0652, 0.1908)	(-0.0331, 0.2307)	(-0.1300, 0.2223)	(-0.0453, 0.3224)	(-0.1375, 0.2104)

Table 7: *The table relates the VIX to the WSB index. VIX is the dependent variable, with ADS, EPU, and lagged VIX as control variables. Values represent coefficients with bootstrapped 95% confidence intervals (10,000 iterations). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.*

The results from the OLS regressions between WSB and VIX can be observed in table 7. A strong and statistically significant positive relationship between WSB and VIX from t - 3 through t + 0 (e.g $\beta = 0.0303$ and $p < 0.01$ at t + 0) can be observed. Increases in WSB are associated with an increase in VIX at the same day and the 3 days prior. However, no significant relationship between WSB and VIX could be observed for future values.

The results in table 8 and 9 suggest a mild contemporaneous positive relationship ($p < 0.1$) between WSB and changes in the HML VOL and Meme VOL portfolios. Increases in WSB can thus be hypothesized to be associated with a relative increase (compared to the general market) in volatility for securities that are frequently mentioned on WallStreetBets, and further that stocks with high market betas experience higher comparative volatility spikes than low beta stocks in days with a high bearish sentiment on WallStreetBets. However, given the high p-values, further analysis would be required to draw any definitive conclusions. Continuing, we notice that there is some evidence in table 8 suggesting a reversal effect in changes in relative volatility for meme stocks following a bearish sentiment day measured through WSB ($\beta = -0.0258,\ p < 0.05$ at $t+1$). When the WSB index is high on a given day, the results thus indicate that volatility (relative to the rest of the market) will decrease for the meme stocks on the following day.

To further investigate the relationship between our WSB index and volatility measures, we conduct Granger Causation tests between our variables. The results in table 10

Predictor	Ret(t-3)	Ret(t-2)	Ret(t-1)	Ret(t+0)	Ret(t+1)	Ret(t+2)	Ret(t+3)	Ret(t+4)	Ret(t+5)
const	0.0032	0.0043	0.0045	0.0050	0.0034	0.0026	0.0025	0.0020	0.0027
	(-0.0173, 0.0224)	(-0.0149, 0.0232)	(-0.0156, 0.0243)	(-0.0161, 0.0249)	(-0.0173, 0.0249)	(-0.0168, 0.0243)	(-0.0184, 0.0228)	(-0.0187, 0.0237)	(-0.0181, 0.0238)
WSB	0.0086	0.0021	0.0061	0.0200*	-0.0258**	-0.0086	-0.0024	-0.0017	0.0029
	(-0.0122, 0.0286)	(-0.0200, 0.0222)	(-0.0173, 0.0321)	(-0.0019, 0.0426)	(-0.0515, -0.0008)	(-0.0352, 0.0162)	(-0.0264, 0.0208)	(-0.0247, 0.0219)	(-0.0222, 0.0272)
ADS	0.0121	0.0114	0.0137	0.0124	0.0066	0.0057	0.0069	0.0099	0.0083
	(-0.0107, 0.0341)	(-0.0125, 0.0353)	(-0.0091, 0.0365)	(-0.0100, 0.0361)	(-0.0205, 0.0305)	(-0.0206, 0.0322)	(-0.0172, 0.0304)	(-0.0147, 0.0328)	(-0.0132, 0.0326)
EPU	0.0089	-0.0135	0.0304***	-0.0346***	0.0169*	-0.0000	-0.0052	0.0083	-0.0002
	(-0.0110, 0.0286)	(-0.0350, 0.0089)	(0.0099, 0.0494)	(-0.0554, -0.0150)	(-0.0013, 0.0360)	(-0.0205, 0.0217)	(-0.0262, 0.0169)	(-0.0154, 0.0324)	(-0.0209, 0.0212)
VIX	0.0022	0.0032	-0.0014	0.0031	-0.0224*	-0.0209*	-0.0129	-0.0117	-0.0123
	(-0.0240, 0.0285)	(-0.0217, 0.0287)	(-0.0247, 0.0253)	(-0.0215, 0.0307)	(-0.0487, 0.0035)	(-0.0458, 0.0026)	(-0.0360, 0.0111)	(-0.0337, 0.0121)	(-0.0361, 0.0114)
Meme VOL lagged 1				-0.3848***	-0.0756**	-0.0211	-0.0270	-0.0652**	-0.0009
				(-0.4558, -0.3132)	(-0.1446, -0.0072)	(-0.0818, 0.0451)	(-0.0801, 0.0303)	(-0.1372, -0.0015)	(-0.0664, 0.0637)
Meme VOL lagged 2			-0.3834***	-0.2266***	-0.0426	-0.0334	-0.0749**	-0.0225	-0.0572*
			(-0.4519, -0.3113)	(-0.2989, -0.1512)	(-0.1096, 0.0249)	(-0.0929, 0.0275)	(-0.1508, -0.0065)	(-0.0870, 0.0430)	(-0.1188, 0.0029)
Meme VOL lagged 3		-0.3818***	-0.2155***	-0.1283***	-0.0425	-0.0735**	-0.0307	-0.0650**	-0.0273
		(-0.4503, -0.3122)	(-0.2881, -0.1506)	(-0.1855, -0.0680)	(-0.1027, 0.0156)	(-0.1416, -0.0093)	(-0.0968, 0.0348)	(-0.1260, -0.0053)	(-0.0908, 0.0338)
Meme VOL lagged 4	-0.3653***	-0.2019***	-0.1093***	-0.0897***	-0.0672*	-0.0233	-0.0719**	-0.0219	-0.0002
	(-0.4310, -0.2993)	(-0.2796, -0.1285)	(-0.1697, -0.0395)	(-0.1437, -0.0334)	(-0.1448, 0.0013)	(-0.0919, 0.0376)	(-0.1277, -0.0132)	(-0.0825, 0.0390)	(-0.0547, 0.0508)
Meme VOL lagged 5	-0.1684***	-0.0896***	-0.0494*	-0.0890***	-0.0053	-0.0594**	-0.0309	0.0180	-0.0553**
	(-0.2386, -0.0984)	(-0.1422, -0.0378)	(-0.1037, 0.0038)	(-0.1532, -0.0243)	(-0.0664, 0.0535)	(-0.1193, -0.0041)	(-0.0900, 0.0259)	(-0.0361, 0.0698)	(-0.1058, -0.0003)

Table 8: *The table relates the change in Meme VOL to the WSB index. The change in Meme VOL is the dependent variable and WSB is the independent variable. Furthermore, ADS, EPU, VIX and 5 day lagging change in Meme VOL are used as control variables. Finally we display bootstrapped 95% confidence intervals with 1000 iterations., *, ** and *** denotes significance at the 10%, 5% and 1% level..*

shows how WSB significantly granger causes Meme VOL for the first three lags. This relationship weakens, for every lag, both in terms of F-values and significance levels (e.g F-value of 8.9149 and $p < 0.01$ for lag 1 and F-value of 4.0811 and $p < 0.05$ for lag 2). However, Meme VOL fails to granger cause WSB at all lags (p-values are consistently above 10%). In table 10, we can also observe how no casual relationship between WSB and the HML VOL can be found. P-values are above 10% and F-values are below 1 for all lags except lag 5 where $F = 1.0339$ and $p > 0.1$. Finally, we can observe how WSB granger causes VIX, with p values below 5% for all lags. The same is true for the test if VIX granger causes WSB, with an exception for lag 1 where $p > 0.1$.

In table 11, one can see how WSB granger causes changes in Meme VOL for the first two lags (e.g $F = 5.0288$ and $p < 0.05$ for lag 1). There are however no evidence for granger causation in the opposite direction. Similar to the results in table 10, there are no granger causation relationship between changes in Meme VOL and WSB in neither direction ($p > 0.1$ for all lags and directions).

Predictor	Ret(t-3)	Ret(t-2)	Ret(t-1)	Ret(t+0)	Ret(t+1)	Ret(t+2)	Ret(t+3)	Ret(t+4)	Ret(t+5)
const	-0.0007	0.0002	0.0004	0.0003	0.0002	-0.0004	-0.0004	-0.0007	-0.0002
	(-0.0232, 0.0215)	(-0.0225, 0.0229)	(-0.0209, 0.0219)	(-0.0208, 0.0222)	(-0.0258, 0.0236)	(-0.0262, 0.0240)	(-0.0256, 0.0234)	(-0.0274, 0.0237)	(-0.0262, 0.0269)
WSB	0.0011	0.0094	-0.0033	0.0208*	-0.0158	0.0036	0.0025	-0.0097	-0.0026
	(-0.0228, 0.0223)	(-0.0122, 0.0308)	(-0.0282, 0.0199)	(-0.0015, 0.0433)	(-0.0431, 0.0119)	(-0.0197, 0.0282)	(-0.0235, 0.0260)	(-0.0341, 0.0169)	(-0.0279, 0.0225)
ADS	0.0009	-0.0060	-0.0035	-0.0083	-0.0064	-0.0073	-0.0081	-0.0069	-0.0071
	(-0.0366, 0.0401)	(-0.0425, 0.0339)	(-0.0406, 0.0355)	(-0.0471, 0.0334)	(-0.0503, 0.0366)	(-0.0437, 0.0334)	(-0.0468, 0.0307)	(-0.0441, 0.0317)	(-0.0399, 0.0273)
EPU	-0.0026	-0.0204*	0.0142	-0.0052	0.0113	0.0056	-0.0163	0.0151	-0.0051
	(-0.0287, 0.0240)	(-0.0434, 0.0035)	(-0.0128, 0.0415)	(-0.0269, 0.0167)	(-0.0117, 0.0352)	(-0.0201, 0.0295)	(-0.0416, 0.0110)	(-0.0108, 0.0389)	(-0.0281, 0.0188)
VIX	0.0129	0.0073	0.0229	0.0163	0.0035	0.0008	-0.0017	0.0021	-0.0033
	(-0.0235, 0.0521)	(-0.0302, 0.0473)	(-0.0118, 0.0554)	(-0.0221, 0.0520)	(-0.0314, 0.0374)	(-0.0317, 0.0348)	(-0.0378, 0.0311)	(-0.0281, 0.0322)	(-0.0363, 0.0294)
HML VOL lagged 1				-0.4960***	-0.1178***	-0.0144	-0.0171	-0.0389	-0.0763**
				(-0.5594, -0.4321)	(-0.1837, -0.0467)	(-0.0732, 0.0509)	(-0.0771, 0.0413)	(-0.0964, 0.0214)	(-0.1384, -0.0160)
HML VOL lagged 2			-0.4875***	-0.3496***	-0.0706*	-0.0257	-0.0481	-0.0850***	0.0012
			(-0.5579, -0.4207)	(-0.4123, -0.2827)	(-0.1386, 0.0042)	(-0.0922, 0.0424)	(-0.1185, 0.0184)	(-0.1498, -0.0203)	(-0.0646, 0.0727)
HML VOL lagged 3		-0.4738***	-0.3342***	-0.2150***	-0.0633*	-0.0581*	-0.0980***	0.0041	0.0291
		(-0.5445, -0.4112)	(-0.4001, -0.2699)	(-0.2772, -0.1496)	(-0.1255, 0.0003)	(-0.1265, 0.0088)	(-0.1630, -0.0301)	(-0.0652, 0.0760)	(-0.0413, 0.1002)
HML VOL lagged 4	-0.4373***	-0.3093***	-0.1917***	-0.1250***	-0.0731**	-0.1095***	-0.0052	0.0461	-0.0137
	(-0.5052, -0.3753)	(-0.3684, -0.2496)	(-0.2511, -0.1256)	(-0.1886, -0.0630)	(-0.1392, -0.0065)	(-0.1748, -0.0442)	(-0.0731, 0.0610)	(-0.0212, 0.1117)	(-0.0852, 0.0546)
HML VOL lagged 5	-0.2389***	-0.1484***	-0.0917***	-0.0675***	-0.1112***	-0.0225	0.0386	0.0167	-0.0716**
	(-0.2882, -0.1866)	(-0.2033, -0.0859)	(-0.1392, -0.0407)	(-0.1236, -0.0115)	(-0.1713, -0.0522)	(-0.0801, 0.0356)	(-0.0209, 0.0968)	(-0.0422, 0.0761)	(-0.1337, -0.0122)

Table 9: *The table relates the change in HML VOL to the WSB. The change in HML VOL is the dependent variable and the WSB is the independent variable. Furthermore, ADS, EPU, VIX and 5 day lagging change HML VOL are used as control variables. Finally we display bootstrapped 95% confidence intervals with 1000 iterations., *, ** and *** denotes significance at the 10%, 5% and 1% level.*

The results in this section thus indicate that the WSB index is strongly correlated with past and current market expectations on future volatility. In periods where the WSB index is high, it is likely that the market has expected volatility to be high in the past days, as indicated through the positive relationship between WSB and prior day VIX values. The fact that the OLS regressions between the WSB index and VIX does not reveal any significant reversal effects, in line with what we observed in the returns section, suggests that markets are not overreacting to sentiment shocks observed on WallStreetBets. However, the results in the granger causation analysis, with both VIX and WSB granger causing each other, paint a different picture and opens up for the possibility that passed values of WSB can be used to forecast future VIX movements. Although, the bidirectional relationship between the variables could also simply suggest that VIX and WSB are heavily correlated and that an unobserved third variable is actually causing both of them. Further analysis would thus be needed to investigate their relationship.

Moving on, in this section we found some evidence that meme stocks volatilities are more susceptible to WSB fluctuations than the general market. Both the OLS regressions for $t + 1$ and granger causality tests [5] showed that past values of WSB can be used to improve forecasting on changes in the meme stocks relative volatilities. The reversal effect observed in the regression results therefore suggests that when WSB is high, meme stocks relative volatility will tend to decrease in the following day.

[5] both with Meme VOL and change in Meme VOL

Lag	Test	WSB → Meme VOL	Meme VOL → WSB	WSB → HML VOL	HML VOL → WSB	WSB → VIX	VIX → WSB
Lag 1	SSR F	8.9149***	2.6395	0.1702	0.9307	5.3735**	0.9646
	Chi²	8.9324***	2.6447	0.1705	0.9325	5.3841**	0.9665
	LR	8.9065***	2.6424	0.1705	0.9323	5.3747**	0.9662
	Param F	8.9149***	2.6395	0.1702	0.9307	5.3735**	0.9646
Lag 2	SSR F	4.0811**	1.2542	0.1336	0.9234	4.0070**	4.6550***
	Chi²	8.1890**	2.5167	0.2681	1.8529	8.0403**	9.3406***
	LR	8.1672**	2.5146	0.2681	1.8518	8.0193**	9.3122***
	Param F	4.0811**	1.2542	0.1336	0.9234	4.0070**	4.6550***
Lag 3	SSR F	2.5409*	0.6765	0.3875	0.5568	3.5024**	4.9374***
	Chi²	7.6578*	2.0387	1.1677	1.6781	10.5555**	14.8802***
	LR	7.6387*	2.0374	1.1673	1.6772	10.5192**	14.8083***
	Param F	2.5409*	0.6765	0.3875	0.5568	3.5024**	4.9374***
Lag 4	SSR F	1.8369	0.4826	0.4286	0.7676	3.3898***	3.9737***
	Chi²	7.3910	1.9417	1.7247	3.0886	13.6395***	15.9888***
	LR	7.3732	1.9404	1.7237	3.0855	13.5790***	15.9058***
	Param F	1.8369	0.4826	0.4286	0.7676	3.3898***	3.9737***
Lag 5	SSR F	1.4474	1.2688	0.3361	1.0339	2.4405**	3.3456***
	Chi²	7.2897	6.3900	1.6927	5.2072	12.2912**	16.8494***
	LR	7.2724	6.3767	1.6917	5.1983	12.2420**	16.7572***
	Param F	1.4474	1.2688	0.3361	1.0339	2.4405**	3.3456***

Table 10: *The table presents the results of Granger Causality tests between the WSB index and the Meme VOL, the HML VOL as well as VIX. For each pair of variables, we conduct bi-directional Granger causality tests over five different lag lengths (Lag 1 to Lag 5), each corresponds to one trading day. The reported statistics include the SSR-based F-test, the Chi-squared statistic, the Likelihood Ratio (LR) test, and the parameter-based F-test. Finally, * , ** and *** denotes significance at the 10%, 5% and 1% level*

5 Conclusion

The analysis of the WSB index and its association with the returns and volatilities of different portfolios thus show that increases in WSB coincides with lower returns for the general market and meme stocks that are frequently discussed on WallStreetBets. Further when WSB is high, relative returns of high beta stocks are low. Our results thus show that investors switch from (to) safe to (from) risky stocks in times when sentiment improves (declines) but no causal relationship has been revealed. Furthermore, compared to the empirical results by the likes of Baker and Wurgler (2006)[1], Da (2015)[7] and Tetlock (2007)[17], we find no evidence of mispricings in reactions to such sentiment changes. Instead the lack of significant reversal effects for returns of the SPY, Meme and HML-risk portfolios, along with the significant relationship between prior returns and WSB, combined with WSB being granger caused by both SPY and Meme portfolio returns, indicates that the WSB sentiment index captures the historical and current performance of the portfolios, and provides limited insights into future performance.

Further, the results in the Volatility section are relatively consistent with our findings in the returns section as the OLS regressions indicate that WSB is significantly correlated with past and current expectations on future volatility, but provides limited insights into future VIX values. The bidirectional VIX-WSB granger causal relationship provides a slightly more nuanced view, however, we find it unlikely (given the OLS regression results in both the volatility and return section) that VIX would be caused or preceded by WSB, instead we propose that the bidirectional relationship could potentially be explained by a third omitted variable.

The results in the Granger causality tests suggests that the WSB index has a predictive power over both the Meme VOL portfolio and changes in the Meme VOL portfolio. This would mean that sentiment on WallStreetBets can help predict the relative volatility of popular meme stocks as compared to the market. Moreover, weak evidence for a reversal is found in this relationship where the relative volatility is decreased in the next day. One possible explanation could be that when the WSB index is high, WallStreetBets users pay closer attention to the frequently mentioned stocks on WallStreetBets and start trading those securities to a larger extent, the sudden influx of retail trading activities on those stocks could then inflate intraday volatility on the day of the sentiment shock. Following days when sentiment settles back, that retail trading pressure would presumably decrease resulting in a reversal in relative volatility compared to the general market. However, the lack of a significant relationship ($p > 0.05$) between WSB and same day change in Meme VOL, as well as high p-values in the Granger causality test puts doubt on this suggestion.

Overall, our empirical results thus coincide with the results presented by the likes of Wang (2006) [18], as most of our empirical results suggest that extracting sentiments based on aggregated textual analysis on discourse on WallStreetBets provides limited insights into future security returns and volatility. Instead of having predicting power, the WSB index is a lagging indicator against the measurable outcomes of the market. However, we are hesitant to draw the conclusion that WallStreetBets user activity can't reveal information that could affect future returns. For example the alleged short squeeze by WallStreetBets on Gamestop indicates that the user activity could have a significant impact on certain stock returns. To further develop the research on WallStreetBets sentiment and stock returns and volatilities, we believe that using LLMs to score all (or a filtered subset of) the 50 million comments and submissions in the dataset on future oriented sentiment could improve the WSB index ability to forecast future returns.

Lag	Test	WSB → Meme VOL	Meme VOL → WSB	WSB → HML VOL	HML VOL → WSB	WSB → VIX	VIX → WSB
Lag 1	SSR F	5.0288**	0.0799	0.8288	1.3115	5.3735**	0.9646
	Chi²	5.0387**	0.0801	0.8304	1.3141	5.3841**	0.9665
	LR	5.0304**	0.0801	0.8302	1.3135	5.3747**	0.9662
	Param F	5.0288**	0.0799	0.8288	1.3115	5.3735**	0.9646
Lag 2	SSR F	2.7790*	0.2754	0.4446	0.3233	4.0070**	4.6550***
	Chi²	5.5762*	0.5525	0.8921	0.6488	8.0403**	9.3406***
	LR	5.5661*	0.5524	0.8918	0.6486	8.0193**	9.3122***
	Param F	2.7790*	0.2754	0.4446	0.3233	4.0070**	4.6550***
Lag 3	SSR F	1.8937	0.1155	0.3712	0.6753	3.5024**	4.9374***
	Chi²	5.7071	0.3480	1.1188	2.0353	10.5555**	14.8802***
	LR	5.6965	0.3479	1.1183	2.0339	10.5192**	14.8083***
	Param F	1.8937	0.1155	0.3712	0.6753	3.5024**	4.9374***
Lag 4	SSR F	1.3716	0.8125	0.4897	1.0448	3.3898***	3.9737***
	Chi²	5.5189	3.2691	1.9702	4.2039	13.6395***	15.9888***
	LR	5.5090	3.2656	1.9689	4.1982	13.5790***	15.9058***
	Param F	1.3716	0.8125	0.4897	1.0448	3.3898***	3.9737***
Lag 5	SSR F	1.0303	0.9094	0.4229	1.1285	2.4405**	3.3456***
	Chi²	5.1891	4.5799	2.1301	5.6836	12.2912**	16.8494***
	LR	5.1803	4.5731	2.1286	5.6730	12.2420**	16.7572***
	Param F	1.0303	0.9094	0.4229	1.1285	2.4405**	3.3456***

Table 11: *The table presents the results of Granger Causality tests between the WSB index and the change in Meme VOL, the change in HML VOL as well as VIX. For each pair of variables, we conduct bi-directional Granger causality tests over five different lag lengths (Lag 1 to Lag 5). each corresponds to one trading day. The reported statistics include the SSR-based F-test, the Chi-squared statistic, the Likelihood Ratio (LR) test, and the parameter-based F-test. Finally, * , ** and *** denotes significance at the 10%, 5% and 1% level.*

References

[1] Malcolm Baker and Jeffrey Wurgler. Investor sentiment and the cross-section of stock returns. The Journal of Finance, 61(4):1645–1680, 2006.

[2] Brad M. Barber and Terrance Odean. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. The Review of Financial Studies, 21(2):785–818, 2008.

[3] Michael W. Brandt, Alon Brav, John R. Graham, and Alok Kumar. The idiosyncratic volatility puzzle: Time trend or speculative episodes? The Review of Financial Studies, 23(2):863–899, 2010.

[4] Gregory W. Brown and Michael T. Cliff. Investor sentiment and the near-term stock market. Journal of Empirical Finance, 11(1):1–27, 2004.

[5] James A. Cookson and Marina Niessner. Why don't we agree? evidence from a social network of investors. The Journal of Finance, 75(1):173–227, 2020.

[6] Corporate Finance Institute. Meme stock, 2024. Accessed March 24, 2025.

[7] Zhi Da, Joseph Engelberg, and Pengjie Gao. The sum of all fears: Investor sentiment and asset prices. The Review of Financial Studies, 28(1):1–32, 2015.

[8] J. Bradford De Long, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. Noise trader risk in financial markets. Journal of Political Economy, 98(4):703–738, 1990.

[9] Luke DeVault, Richard Sias, and Laura T. Starks. Sentiment metrics and investor demand. The Journal of Finance, 74(2):985–1024, 2019.

[10] Diego Garcia. Sentiment during recessions. The Journal of Finance, 68(3):1267–1300, 2013.

[11] C. W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3):424–438, 1969.

[12] Roger Koppl. Retrospectives: Animal spirits. Journal of Economic Perspectives, 5(3):203–210, 1991.

[13] Wayne Y. Lee, Christine X. Jiang, and Daniel C. Indro. Stock market volatility, excess returns, and the role of investor sentiment. Journal of Banking & Finance, 26(12):2277–2299, 2002.

[14] Quiver Quantitative. Meme stock rankings, 2024. Accessed March 26, 2025.

[15] David P. Simon and Roy A. Wiggins III. S&p futures returns and contrary sentiment indicators. Journal of Futures Markets, 21(5):447–462, 2001.

[16] Robert F. Stambaugh, Jianfeng Yu, and Yu Yuan. The short of it: Investor sentiment and anomalies. Journal of Financial Economics, 104(2):288–302, 2012.

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

[18] Yaw-Huei Wang, Aneel Keswani, and Stephen J. Taylor. The relationships between sentiment, returns and volatility. International Journal of Forecasting, 22(1):109–123, 2006.

A Use of AI in Thesis

In the course of conducting the empirical analysis AI, namely Chat GPT, has been used to assist with coding, debugging and explaining code and econometric methods. Chat GPT has further been used to check for spelling and grammatical errors as well as formatting LaTeX tables in the thesis.

B Supplementary Tables

Table 12: List of Finance-Related Words Used for Filtering

Finance-Related Words		
buy	sell	calls
puts	bull	bear
SPY	market	stocks
shares	trading	investing
portfolio	dividends	earnings
profits	losses	capital
hedge	short	long
margin	volatility	options
ETF	index	funds
valuation	growth	crash
rally	breakout	FOMO
uptrend	rebound	

Table 13: Bullish and Bearish Words Used for Filtering

Bullish Words		Bearish Words	
bullish	pump	bearish	crash
moon	FOMO	plummet	recession
breakout	uptrend	sell-off	bankruptcy
gains	rally	correction	downtrend
ATH	rocketing	freefall	implosion
exploding	squeezing	meltdown	pain
printing	skyrocketing	nuked	capitulation
sending	all-time high	wrecked	slaughter
parabolic	LFG	blood red	sell pressure
green candles	sending it	profit taking	draining
blast off	loading up	resistance rejected	distribution phase
strong support	buy pressure	market turmoil	pain incoming
going up	next leg up	weak bounce	rugged
moonshot	no resistance	dump it	no support
gap up	momentum building	breaking down	weakness showing
rocket fuel	buy the dip	down only	not looking good
reversal incoming	breakout confirmed	dump confirmed	washed out
bounce back	buy zone	panic selling	floodgates open
accumulation phase	supercycle	slow bleed	sinking
melting up	bull run	exit liquidity	trend reversal confirmed
price discovery	higher highs	ugly chart	bear trap
clear skies ahead	hyper bullish	short the rip	short squeeze incoming
dip bought	big money buying	bagholder	lower highs
stacking	liquidity injection	death cross	liquidity drained
next stop the moon	volatility squeeze	oversupply	sell zone
up only	gap and go	distribution	dump fest
max pain	unloading	bearish engulfing	fakeout
hitting resistance	double top	dead cat bounce	bearish divergence
market manipulation	mass liquidations	big red candle	supply zone
hard rejection	margin calls	overleveraged	total collapse

Lag	Test	WSB → SPY	SPY → WSB	WSB → Meme	Meme → WSB	WSB → HML-Risk	HML-Risk → WSB
Lag 1	SSR F	0.0686	4.1270**	0.2482	6.0778**	0.1224	0.8472
	Chi²	0.0687	4.1351**	0.2487	6.0897**	0.1226	0.8488
	LR	0.0687	4.1295**	0.2486	6.0776**	0.1226	0.8486
	Param F	0.0686	4.1270**	0.2482	6.0778**	0.1224	0.8472
Lag 2	SSR F	0.4478	4.5643**	0.7206	5.6617***	0.7388	1.7777
	Chi²	0.8985	9.1586**	1.4458	11.3606***	1.4824	3.5671
	LR	0.8983	9.1313**	1.4451	11.3186***	1.4817	3.5630
	Param F	0.4478	4.5643**	0.7206	5.6617***	0.7388	1.7777
Lag 3	SSR F	0.4458	4.2474***	1.0866	5.1119***	0.6606	1.6178
	Chi²	1.3436	12.8007***	3.2748	15.4063***	1.9909	4.8757
	LR	1.3430	12.7475***	3.2713	15.3293***	1.9896	4.8680
	Param F	0.4458	4.2474***	1.0866	5.1119***	0.6606	1.6178
Lag 4	SSR F	0.2085	3.7421***	0.8130	4.0171***	0.5531	2.0770*
	Chi²	0.8390	15.0572***	3.2711	16.1635***	2.2256	8.3571*
	LR	0.8388	14.9835***	3.2676	16.0787***	2.2240	8.3344*
	Param F	0.2085	3.7421***	0.8130	4.0171***	0.5531	2.0770*
Lag 5	SSR F	0.1675	3.3334***	0.7023	3.5422***	0.4220	2.1937*
	Chi²	0.8434	16.7879***	3.5369	17.8393***	2.1252	11.0478*
	LR	0.8431	16.6964***	3.5328	17.7360***	2.1237	11.0081*
	Param F	0.1675	3.3334***	0.7023	3.5422***	0.4220	2.1937*
Lag 6	SSR F	0.1519	2.9870***	1.0434	2.9682***	0.4961	2.0486*
	Chi²	0.9195	18.0758***	6.3142	17.9619***	3.0019	12.3974*
	LR	0.9192	17.9697***	6.3011	17.8571***	2.9989	12.3473*
	Param F	0.1519	2.9870***	1.0434	2.9682***	0.4961	2.0486*
Lag 7	SSR F	0.2033	2.5798**	1.1262	2.6733***	0.9743	2.2597*
	Chi²	1.4372	18.2380**	7.9617	18.8986***	6.8878	15.9746*
	LR	1.4365	18.1299**	7.9410	18.7826***	6.8723	15.8916*
	Param F	0.2033	2.5798**	1.1262	2.6733***	0.9743	2.2597*
Lag 8	SSR F	0.1805	2.5249***	1.0070	2.3296**	0.9625	2.0437*
	Chi²	1.4600	20.4267***	8.1472	18.8466**	7.7872	16.5338*
	LR	1.4593	20.2911***	8.1255	18.7311**	7.7674	16.4448*
	Param F	0.1805	2.5249***	1.0070	2.3296**	0.9625	2.0437*
Lag 9	SSR F	0.1970	2.4331***	0.8617	2.1360**	0.8488	1.8498*
	Chi²	1.7956	22.1740***	7.8535	19.4667**	7.7359	16.8587*
	LR	1.7946	22.0142***	7.8334	19.3435**	7.7164	16.7661*
	Param F	0.1970	2.4331***	0.8617	2.1360**	0.8488	1.8498*
Lag 10	SSR F	0.2564	2.4628***	0.8585	1.9871**	0.8475	1.7388*
	Chi²	2.6003	24.9724***	8.7053	20.1485**	8.5932	17.6315*
	LR	2.5981	24.7698***	8.6805	20.0164**	8.5691	17.5303*
	Param F	0.2564	2.4628***	0.8585	1.9871**	0.8475	1.7388*

Table 14: *The table presents the results of bi-directional Granger Causality tests between WSB, SPY, Meme, and HML-risk for lags 1-10. The table reports the SSR-based F-test, Chi-squared statistic, Likelihood Ratio (LR) test, and parameter-based F-test. Finally, * , ** and *** denotes significance at the 10%, 5% and 1% level.*

Lag	Test	WSB → SPY	SPY → WSB	WSB → Meme	Meme → WSB	WSB → HML-Risk	HML-Risk → WSB
Lag 11	SSR F	0.2574	2.2692***	0.7775	1.8765**	0.7606	1.5739*
	Chi²	2.8745	25.3447***	8.6838	20.9582**	8.4948	17.5781*
	LR	2.8718	25.1360***	8.6591	20.8152**	8.4712	17.4773*
	Param F	0.2574	2.2692***	0.7775	1.8765**	0.7606	1.5739*
Lag 12	SSR F	0.2385	2.0440**	0.7096	1.7888**	0.7532	1.4934*
	Chi²	2.9095	24.9384**	8.6571	21.8246**	9.1900	18.2207*
	LR	2.9067	24.7362**	8.6326	21.6695**	9.1623	18.1124*
	Param F	0.2385	2.0440**	0.7096	1.7888**	0.7532	1.4934*
Lag 13	SSR F	0.2440	1.8596**	0.8089	1.7768**	0.7121	1.3676*
	Chi²	3.2299	24.6124**	10.7058	23.5157**	9.4246	18.1006*
	LR	3.2265	24.4153**	10.6683	23.3357**	9.3955	17.9936*
	Param F	0.2440	1.8596**	0.8089	1.7768**	0.7121	1.3676*
Lag 14	SSR F	0.2449	1.7102**	0.8166	1.6675*	0.6836	1.2782
	Chi²	3.4948	24.4089**	11.6547	23.7989**	9.7562	18.2431
	LR	3.4908	24.2149**	11.6102	23.6143*	9.7251	18.1345
	Param F	0.2449	1.7102**	0.8166	1.6675*	0.6836	1.2782
Lag 15	SSR F	0.2953	1.7050**	0.8039	1.7198**	0.6877	1.1535
	Chi²	4.5214	26.1088**	12.3097	26.3355**	10.5308	17.6627
	LR	4.5147	25.8868**	12.2600	26.1096**	10.4944	17.5607
	Param F	0.2953	1.7050**	0.8039	1.7198**	0.6877	1.1535
Lag 16	SSR F	0.2751	1.5546*	0.7868	1.6317*	0.7532	1.1651
	Chi²	4.4988	25.4275*	12.8692	26.6876**	9.1900	19.0562
	LR	4.4922	25.2168*	12.8150	26.4555**	9.1623	18.9375
	Param F	0.2751	1.5546*	0.7868	1.6317*	0.7532	1.1651
Lag 17	SSR F	0.4127	1.4643*	0.7601	1.5208*	0.7601	0.6859
	Chi²	7.1813	25.4820*	13.2269	26.4639*	11.9353	18.8956
	LR	7.1643	25.2702*	13.1695	26.2356*	11.8886	18.7788
	Param F	0.4127	1.4643*	0.7601	1.5208*	0.7601	0.6859
Lag 18	SSR F	0.3928	1.5150*	0.7682	1.4589*	0.6683	0.8449
	Chi²	7.2467	27.9528*	14.1736	26.9171*	12.3305	19.2220
	LR	7.2294	27.6980*	14.1077	26.6808*	12.2806	19.1011
	Param F	0.3928	1.5150*	0.7682	1.4589*	0.6683	0.8449
Lag 19	SSR F	0.5554	1.3918	0.7753	1.3728	0.7753	0.7617
	Chi²	10.8325	27.1430	15.1209	26.7735	12.2458	19.1439
	LR	10.7939	26.9025	15.0459	26.5396	12.1966	19.1129
	Param F	0.5554	1.3918	0.7753	1.3728	0.7753	0.7617
Lag 20	SSR F	0.5564	1.3307	0.7316	1.2969	0.8475	0.5996
	Chi²	11.4372	27.3559	15.0403	26.6605	8.5932	12.3266
	LR	11.3942	27.1115	14.9661	26.4284	8.5691	12.2767
	Param F	0.5564	1.3307	0.7316	1.2969	0.8475	0.5996

Table 15: *The table presents the results of bi-directional Granger Causality tests between WSB, SPY, Meme, and HML-risk for lags 11-20. The table reports the SSR-based F-test, Chi-squared statistic, Likelihood Ratio (LR) test, and parameter-based F-test. Finally, * , ** and *** denotes significance at the 10%, 5% and 1% level*

Predictor	Ret(t-3)	Ret(t-2)	Ret(t-1)	Ret(t+0)	Ret(t+1)	Ret(t+2)	Ret(t+3)	Ret(t+4)	Ret(t+5)
const	0.6957***	0.5730***	0.5545***	0.5277***	0.7324***	0.8322***	0.9108***	0.9973***	1.0609***
	(0.5599, 0.8148)	(0.4377, 0.7119)	(0.4133, 0.6907)	(0.3886, 0.6687)	(0.5773, 0.8944)	(0.6587, 1.0078)	(0.7328, 1.0846)	(0.8182, 1.1658)	(0.8866, 1.2341)
WSB	-0.0002	0.0090	-0.0028	0.0214**	0.0046	0.0080	0.0114	0.0019	-0.0017
	(-0.0239, 0.0208)	(-0.0127, 0.0309)	(-0.0269, 0.0193)	(0.0001, 0.0430)	(-0.0214, 0.0301)	(-0.0168, 0.0352)	(-0.0148, 0.0376)	(-0.0245, 0.0286)	(-0.0302, 0.0268)
ADS	0.0252	0.0140	0.0155	0.0096	0.0103	0.0065	0.0010	-0.0029	-0.0077
	(-0.0095, 0.0638)	(-0.0204, 0.0535)	(-0.0199, 0.0548)	(-0.0277, 0.0492)	(-0.0306, 0.0490)	(-0.0346, 0.0504)	(-0.0414, 0.0448)	(-0.0447, 0.0402)	(-0.0458, 0.0337)
EPU	-0.0069	-0.0231**	0.0117	-0.0067	0.0026	0.0076	-0.0082	0.0071	0.0010
	(-0.0325, 0.0190)	(-0.0444, -0.0006)	(-0.0126, 0.0369)	(-0.0283, 0.0155)	(-0.0216, 0.0264)	(-0.0163, 0.0321)	(-0.0347, 0.0185)	(-0.0213, 0.0356)	(-0.0282, 0.0269)
VIX	0.0766***	0.0606***	0.0752***	0.0673***	0.0893***	0.0995***	0.1061***	0.1170***	0.1189***
	(0.0356, 0.1194)	(0.0184, 0.1067)	(0.0395, 0.1126)	(0.0230, 0.1071)	(0.0399, 0.1385)	(0.0493, 0.1497)	(0.0502, 0.1556)	(0.0667, 0.1621)	(0.0680, 0.1615)
HML VOL lagged 1				0.4639***	0.3383***	0.3174***	0.2897***	0.2413***	0.1661***
				(0.4009, 0.5246)	(0.2682, 0.4075)	(0.2457, 0.3853)	(0.2199, 0.3602)	(0.1705, 0.3137)	(0.0940, 0.2398)
HML VOL lagged 2			0.4741***	0.1280***	0.1769***	0.1628***	0.1241***	0.0723*	0.1540***
			(0.4070, 0.5379)	(0.0563, 0.2046)	(0.1008, 0.2516)	(0.0869, 0.2428)	(0.0473, 0.1979)	(-0.0021, 0.1398)	(0.0767, 0.2378)
HML VOL lagged 3		0.4944***	0.1318***	0.1177***	0.1301***	0.0958**	0.0367	0.1193***	0.1528***
		(0.4278, 0.5540)	(0.0616, 0.2046)	(0.0546, 0.1778)	(0.0615, 0.1983)	(0.0227, 0.1659)	(-0.0368, 0.1154)	(0.0406, 0.2034)	(0.0738, 0.2321)
HML VOL lagged 4	0.5390***	0.1487***	0.1242***	0.0756**	0.0732**	0.0208	0.1039**	0.1386***	0.1049**
	(0.4717, 0.5964)	(0.0800, 0.2202)	(0.0644, 0.1842)	(0.0099, 0.1390)	(0.0081, 0.1405)	(-0.0629, 0.0984)	(0.0162, 0.1894)	(0.0596, 0.2129)	(0.0246, 0.1811)
HML VOL lagged 5	0.2491***	0.1826***	0.1013***	0.0542**	0.0587*	0.1501***	0.1685***	0.1248***	0.0990***
	(0.1886, 0.3166)	(0.1283, 0.2347)	(0.0429, 0.1582)	(0.0028, 0.1070)	(-0.0034, 0.1213)	(0.0816, 0.2232)	(0.1005, 0.2400)	(0.0604, 0.1973)	(0.0284, 0.1715)

Table 16: *The table relates the HML VOL to the WSB index. The HML VOL is the dependent variable and the WSB index is the independent variable. ADS, EPU, VIX, and five lagged values of HML VOL are used as controls. 95% bootstrapped confidence intervals (1000 iterations) in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.*

Predictor	Ret(t-3)	Ret(t-2)	Ret(t-1)	Ret(t+0)	Ret(t+1)	Ret(t+2)	Ret(t+3)	Ret(t+4)	Ret(t+5)
const	0.2875	0.2517	0.2087	0.2146	0.2792	0.3262	0.3730	0.4039	0.4233
	(0.1920, 0.3799***)	(0.1560, 0.3443***)	(0.1192, 0.2976***)	(0.1250, 0.3108***)	(0.1774, 0.3858***)	(0.2113, 0.4396***)	(0.2599, 0.4859***)	(0.2805, 0.5290***)	(0.2988, 0.5510***)
WSB	0.0016	-0.0040	0.0011	0.0149	-0.0124	-0.0220	-0.0253	-0.0277	-0.0252
	(-0.0194, 0.0221)	(-0.0259, 0.0162)	(-0.0219, 0.0269)	(-0.0070, 0.0374)	(-0.0379, 0.0137)	(-0.0480, 0.0052)	(-0.0535, 0.0004)*	(-0.0569, 0.0023)*	(-0.0578, 0.0089)
ADS	0.0053	0.0061	0.0096	0.0081	0.0136	0.0179	0.0237	0.0332	0.0407
	(-0.0187, 0.0287)	(-0.0180, 0.0310)	(-0.0138, 0.0343)	(-0.0153, 0.0316)	(-0.0106, 0.0395)	(-0.0087, 0.0454)	(-0.0031, 0.0492)*	(0.0050, 0.0604)**	(0.0118, 0.0723***)
EPU	0.0087	-0.0139	0.0298	-0.0347	-0.0179	-0.0176	-0.0226	-0.0144	-0.0143
	(-0.0110, 0.0285)	(-0.0359, 0.0076)	(0.0094, 0.0486***)	(-0.0550, -0.0147***)	(-0.0403, 0.0038)	(-0.0440, 0.0076)	(-0.0497, 0.0035)*	(-0.0437, 0.0132)	(-0.0436, 0.0123)
VIX	0.0326	0.0297	0.0206	0.0255	0.0097	-0.0064	-0.0145	-0.0230	-0.0336
	(0.0038, 0.0616)**	(0.0005, 0.0591)**	(-0.0043, 0.0510)	(-0.0041, 0.0574)*	(-0.0222, 0.0436)	(-0.0415, 0.0288)	(-0.0472, 0.0250)	(-0.0573, 0.0175)	(-0.0694, 0.0016)*
Meme VOL lagged 1				0.6051	0.5256	0.5042	0.4759	0.4077	0.4084
				(0.5351, 0.6784***)	(0.4503, 0.5977***)	(0.4296, 0.5788***)	(0.3919, 0.5546***)	(0.3166, 0.4953***)	(0.3157, 0.4984***)
Meme VOL lagged 2			0.6038	0.1547	0.1855	0.1752	0.1273	0.1678	0.1141
			(0.5350, 0.6757***)	(0.0901, 0.2246***)	(0.1089, 0.2673***)	(0.0872, 0.2559***)	(0.0401, 0.2199***)	(0.0701, 0.2600***)	(0.0165, 0.2123**)
Meme VOL lagged 3		0.6090	0.1642	0.1008	0.0992	0.0637	0.1097	0.0645	0.0997
		(0.5405, 0.6810***)	(0.0945, 0.2324***)	(0.0281, 0.1717***)	(0.0182, 0.1774**)	(-0.0210, 0.1495)	(0.0062, 0.2086)**	(-0.0363, 0.1687)	(-0.0113, 0.2107*)
Meme VOL lagged 4	0.6506	0.1821	0.1058	0.0462	0.0201	0.0790	0.0414	0.0801	0.1160
	(0.5890, 0.7118***)	(0.1122, 0.2460***)	(0.0348, 0.1795***)	(-0.0154, 0.1056)	(-0.0536, 0.0972)	(-0.0069, 0.1657)*	(-0.0534, 0.1437)	(-0.0198, 0.1827)	(0.0104, 0.2170**)
Meme VOL lagged 5	0.2718	0.1414	0.0706	0.0360	0.0958	0.0919	0.1477	0.1740	0.1514
	(0.2159, 0.3297***)	(0.0695, 0.2157***)	(0.0102, 0.1248**)	(-0.0167, 0.0842)	(0.0265, 0.1634***)	(0.0180, 0.1619**)	(0.0677, 0.2190***)	(0.0820, 0.2529***)	(0.0639, 0.2347***)

Table 17: *The table relates the Meme VOL to the WSB index. The Meme VOL is the dependent variable and the WSB index is the independent variable. Furthermore, ADS, EPU, VIX and 5 day lagging Meme VOL are used as control variables. Bootstrapped 95% confidence intervals with 1000 iterations are shown. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively.*