Contents

1 Introduction **2**

2 Literature Review **2**

2.1 Noise Trader Theory 2

2.2 Individual Investor Sentiment, News & Social Media 4

2.3 News & Social Media Sentiment 6

2.4 Wavelet Analysis 7

2.5 GameStop Related Literature 8

2.6 Hypotheses & Contribution 9

3 Methodology **9**

3.1 Data 9

4 Results **15**

4.1 Text Analysis 15

4.2 Wavelet Coherence 18

4.3 Comovements 23

5 Discussion **28**

Conclusion **28**

References **34**

Appendix **36**

Note: the master thesis code repository can be found here `https://github.com/kevintrng/MasThe/sentiment`

1 Introduction

2 Literature Review

Following our motivation we categorise the literature survey into five general parts: noise trader theory, investor sentiment, news and social media sentiment, market manipulation and short-selling and lastly, wavelet analysis. In addition to these strands of literature, we also separately discuss numerous recent academic works which similarly cover the GameStop retail trading saga, as it gained plenty of attention from both investors, outsiders and academia alike.

2.1 Noise Trader Theory

Given the thesis' focus on individual investors and investor sentiment, and the overall nature of the GameStop trading frenzy, the following section motivates our departure from the efficient market hypothesis, the overview of which can be found in the seminal work of Fama (1970) [1]. The more recent survey of theory is given by Yen and Lee (2008) [2]. Related to the emphasis of the role of irrational agents, trading on information and noise, and imperfect, respectively limited, arbitrage, we particularly present the strand of literature within *noise trader theory*. Thus, allowing the possibility of long-term deviations from the fundamentals.[1]

Although, firstly, there have been numerous stylised factors, respectively financial anomalies, backed by extensive empirical evidence irreconcilable with the theory of efficient markets, some of which we discuss to motivate our departure. Including, but not limited to, Shiller's proof of excess volatility [3], Bond and Thaler's (1985) overshooting [4], day-of-the-week effect among others. Further, the "January effect" or "Turn-of-the-year effect", first documented by Rozeff and Kinney (1976), [5] describes the empirical finding that small stocks generate significantly higher returns than their large counterparts and market indices in January, especially following a decline in December. Roll (1983) and Reinganum (1983) [6, 7] ascribe the effect to the tax-loss-selling, creating price pressures on stocks that have poorly performed throughout the year, after the turn-of-the-year this pressure is alleviated and returns revert back to the equilibrium. Their hypothesis is further generalised by Ritter (1988) [8] which emphasises the role of buying and selling behaviour of individual investors, particularly the reinvesting in January following the

[1] redundant?

tax-loss-motivated selling before the end of the year. The "January effect" is one of the many market inefficiencies documenting the price effect driven by factors other than fundamental news. In theory, given the existence of non-constrained arbitrage, it could not exist as temporary price trading patterns would be eliminated.

Following Kyle's (1985) [9] definition of random noise trader, Black (1986) [10] establishes the *noise trader theory* representing a distinct departure from the theory of efficient markets. Black [10] divides traders into information and noise traders, with the latter trading on noise as if it were information, but ultimately both groups do not know for certain whether they are trading on information or noise and whether or not has the information already been priced in. Moreover, two reasons for the prevalence of noise trading are provided. First, investors simply like to trade on noise and second, they believe they are trading on information.

In reaction to the documented empirical shortfalls of the efficient market hypothesis, Shleifer and Summers (1990) [11] argue in favour of pursuing an alternative approach, one based on Black's [10] model, motivated by two central themes. First, the existence of limited arbitrage, due to its inherent riskiness which extends beyond fundamental risk and also includes the probability that the mispricing becomes even more extreme in the future, described as *future resale price risk*. Second, the existence of irrational agents whose demand for risky assets is driven by their beliefs and sentiment irrespective of fundamentals. De Long, Shleifer, Summers and Waldmann (hereafter DSSW) in a series of works (1989, 1990a, 1990b) [12, 13, 14] further expand on the *future resale price risk* and the unpredictability of investor sentiment, described as the *noise trader risk*, and develop a model in which sophisticated investors and noise traders interact, where the latter trade on incorrect beliefs and pseudo-signals. In addition to being able to explain some of the financial anomalies such as Mehra and Prescott (1985) [15] equity premium puzzle or undervaluation of closed-end mutual funds[2], it is shown that noise traders can earn higher expected returns from bearing the noise trader risk they create.

Nonetheless, as noted above, it is not of our interest to directly address the efficiency of the markets or test the efficient market hypothesis. Rather, we use the alternative to the efficient market hypothesis, the noise trader theory to facilitate our analysis and provide context. By documenting the GameStop retail trading frenzy we aim to empirically show the fundamental notions of the theory in works, using the framework of investor sentiment, irrational beliefs and limited arbitrage. Provided the GameStop saga was primarily driven by individual retail investors, our thesis relates to this particular strand of literature examining the effect of individual investors and their sentiment.

[2]missing citation and explanation of the issue

2.2 Individual Investor Sentiment, News & Social Media

Included in the section above the paper series of De Long et al. [12, 13, 14] in general indicate that the investor sentiment generates systematic risk possibly leading to large deviations from the fundamental values which moreover, can sustain even in the long-term with the help of limited arbitrage. Following these studies' conclusion regarding the closed-end funds (CEF) pricing, Lee et al. (1991) and Brown (1999) [16, 17] empirically validate the significant relationship between investor sentiment, respectively noise trader risk, and CEF puzzle. In addition, the latter shows a significant impact on CEF's volatility. Further evidence consistent with the noise trader theory's conclusions is provided by Bodurtha et al. (1995) suggesting that closed-end country fund premiums may be related to the U.S. market sentiment, Pontiff's (1997) proof of CEF's excess volatility or the parallel between CEF and the U.K. real estate companies in Barkham et al. (1999) [18, 19, 20]. The effect of investor sentiment using various proxies, including, but not limited to, equity share offerings, closed-fund discounts, net mutual fund redemptions or odd-lot sales, on the cross-section of stock returns is largely documented, see Swaminathan (1996), Kothari and Shanken (1997), Neal and Wheatley (1998) or Baker and Wurgler (2000, 2006) [21, 22, 23, 24, 25]. A short overview of the literature attempting to link individual investor sentiment and the stock market in Baker and Wurgler (2007) [26]. Moreover, they demonstrate the predictive power of a designed six-part measure of sentiment from the review of literature, which is composed of closed-end fund discount, NYSE share turnover, equity share in new issues, number of IPOs, and first-day IPO returns.

The DSSW model [12, 13, 14] predicts that, as sophisticated and noise traders interact on the market, prices can largely deviate from the fundamental values with the latter trading on pseudo-signals, respectively non-fundamental news. Related, and considering the spotlight on the individual, respectively retail, investors by the GameStop saga, we move to sentiment measures more closely related to investor psychology. Specifically, following up is the strand of literature more intimately examining the effect of the sentiment using micro-level data, including brokerage data on individual transactions, trade imbalances, etc. In addition, the survey of the literature studying in detail the retail investors' behavioural tendencies is also included. Consistent with the noise trader theory, the presented literature ought to primarily challenge the prevailing notions on the impact of individual investors, particularly related to their ability to affect prices, coordinated behaviour among a large group of retail investors and lastly, the belief that their coordinated action ought to be cancelled out by rational arbitrageurs and thus, negating their effect on prices. Note that due to the voluminous literature documenting the effect of sentiment on stock markets, we are inherently selective and merely attempt to present the most relevant studies emphasising the role of individual investors and their behavioural tendencies.[3].

There exists extensive evidence that retail investors may excessively respond to pseudo-signals as documented by their suboptimal, respectively excessive, trading activity, mainly caused by overconfidence and detrimental to their expected returns, which is not explained by the theories of rational investors, as shown by Lee (1992), Odean (1999), Barber and Odean (2000, 2001) or Chang, Hsieh and Wang (2015) [27, 28, 29, 30, 31]. Notably, Barber and Odean (2002) [32] observe retail investors' increased traded activity, particularly of speculative type, when switching to online trading in the early 1990s and hypothesise that the concomitant worsening performance is attributable to cognitive biases reinforcing overconfidence such as self-attribution or illusion of control produced by the enhanced access to technology and information.

Beyond the specific behavioural tendencies, there is substantial support for the theory that individual investors act in concert displaying similar trading behaviour across different categories of stocks, as documented by Jackson (2003) [33] for Australian investors, Feng and Seasholes (2004) [34] using Chinese-based dataset and others [35]. Separately, Barber and Odean (2008) [36], using abnormal 1-day volume, previous 1-day return and news mention as proxies for attention, demonstrate that the retail investors' buying decisions, in particular, are largely driven by the attention drawn stock on a given day. Kumar and Lee (2006) [37] in addition to identifying common tendencies of individual investors' trading activity[4], shows that retail sentiment, measured by the buy-sell imbalance, is a significant predictor of returns, particularly of firms with high retail concentration characterised by small capitalisation, high book-to-market ratio, low price and lower institutional ownership). Moreover, relating to the noise trader theory's limited arbitrage argument, higher arbitrage costs is linked to higher sensitivity to retail sentiment. Similarly, Barber, Odean and Zhu (2009) [38] find that retail trade imbalances predict returns over the short horizon, although the effect in the longer horizons is limited to difficult to arbitrage stocks, i.e. high idiosyncratic volatility, small capitalisation stocks, similar evidence is produced by Kaniel, Saar and Titman (2008) [39].

While the above literature illustrates the effect of attention on retail investors' trading activity and the susceptibility of highly volatile stocks to sentiment, others have attempted to reverse the optics and found a significant positive relationship between retail traded stocks and volatility, see Foucault, Sraer and Thesmar (2011) [40], Andrade, Chang,

[3] perhaps refer to some overview. Some overview can be found in [26]

[4] It is shown that across a different group of stocks, retail investors tend to buy particular stocks when selling other and vice versa. Moreover, the behaviour is replicated independently of other investors, i.e. if one group of investors buys (sells) other groups independently also buy (sell).

Seasholes (2008) [41] or Brandt et al. (2010) [42].

Concluding this section related to the role of investor sentiment and behavioural tendencies on a lighter note, some of the related fringe literature is presented. Following works immerse deeper into the investor psychology and instead of sentiment attempt to measure mood, using both exogenous and endogenous factors to emotions. Kamstra, Kramer and Levi (2003) [43] link seasonal affective disorder (SAD) with stock market returns, explaining the relatively lower returns during autumn and relatively higher returns after the winter solstice, with the effect of SAD particularly strong for northern countries. Another discrete measure is offered by Edmans, García and Norli (2007) [44], who demonstrate that loss in major football fixtures predicts next-day significant negative returns. Goetzmann et al. (2015) [45] find that weather-based indicators (such as cloud cover) impact institutional investors' decisions. Yet another recent study piquing interest is Edmans et al. (2021) [46] using music sentiment, derived from Spotify's 200 top songs in a given country, as a proxy of investor mood. Corresponding to the evidence from the above-mentioned literature, music sentiment was able to (positively) predict same-period returns and subsequent price reversal. Moreover, it was also able to predict mutual fund flows and, in the absolute form, volatility.

2.3 News & Social Media Sentiment

So far we have emphasised both the economic fundamentals linked and endogenous measures, related to the latter following strand of literature tackles the transmission mechanism, in particular the media of information be it financial reports, news or social media. The expeditious development in communications technology encompassing the vast ocean of information and importantly the increased access to it, has led to advancement of various methods trying to capture the signalling quality of the information, some of which are central to our thesis. Although, the study of the impact of news is not novel, see Cutler, Poterba and Summers (1989) [47] who however fail to find a significant link. between market aggregate moves and news coverage, Gidofalvi and Elkan's (2001) [48] simple classification algorithm predicting price movements based on news articles, or similar attempt at predicting the impact of news stories using text classification based on SVMs by Fung, Yu and Lu (2005) [49]. Further, text mining techniques have been largely used to predict stock returns, turnover or volatility, ranging from assessing similarity of news (Tetlock, 2007) [50], bag of word or named entity recognition (Schumaker and Chen, 2009) [51].

Moving beyond the realm of traditional news media, alternative sources of sentiment have also piqued large interest from the academic literature, as a source of sentiment that arises from the expressions of and interactions between the individual investors. Online search trends or wikipedia page views have been shown as a significant indicators of stock returns of individual stocks, see Da, Engelberg and Gao (2011, 2015) [52, 53], Preis, Moat and Stanley (2013) [54], or Moat et al. (2013) [55]. Yet another interesting territory, and highly relevant for our thesis, is social media. Antweiler and Frank (2004) [56] show internet message board activity, Yahoo! Finance and Raging Bull in particular, predict stock returns, turnover and volatility. Similarly, using popular stock trading social media Seeking Alpha, Chen et al. (2014) [57] predict future stock returns and earnings surprises. Indeed, even Reddit's r/wallstreetbets subreddit's so.called due dilligence reports have found to significantly predict stock returns in the past as shown by Bradley et al. (2021) [58].

Lastly for this section, we would like to highlight some of the interesting findings pertinent to our analysis. Firstly, the challenging of the uniform nature of investor sentiment derived from various sources. Jiao, Veiga and Walther (2020) differentiate between social media and news sentiment, in particular showing that social media more resemble *echo chambers* where existing fundamental signals are only mulled over, suggesting that high social media coverage is followed by periods of high volatility and turnover compared to the opposite effect of news sentiment. Secondly, consistent with the findings of Antweiler and Frank (2004) [56] Atkins, Niranjan and Gerding (2018) [59] find, using intraday data and Reuters news archive, that sentiment predicts volatility better rather than stock returns.

2.4 Wavelet Analysis

Following section briefly presents literature related to our central method, wavelet analysis. Wavelets have found their use across vast number of disciplines ranging from meteorology, oceanography, geology to signal and image processing (Kumar and Foufoula-Georgiou, 1997 [60]; Kronland-Martinet, 1988 [61]). Notably, the field of economics and finance have been a rather late addition to list of disciplines utilising the wavelet methods. Some interesting applications include Ramsey and Zhang's (1996) study of foreign exchange data structure, or Ramsey and Lampart's (1998) [62] examination of money-income relationship, for an overview refer to Crowley (2007) [63]. Regarding the finance related works, wavelet analysis has been leveraged predominantly in studying comovements (Rua and Nunes, 2009) [64] financial contagion [5]. For a more detailed introduction to the wavelet methods we refer to Torrence and Compo (1998) [65] or previously mentoned Crowley (2007) [63].

first original applications

[5] missing citation

economics and finance favouring the method recently

comovements important for our analysis and hint at financial contagion and gamestop related applications

2.5 GameStop Related Literature

The last literature survey concerns the recent academic works exploring the GameStop saga from various angles. The examination of recent findings and hypotheses serves both for validation purposes and establishing a common base for the literature's findings and thesis' goals.

In the first of a series of papers Umar et al. (2021a, 2021b, 2021c) [66, 67, 68] using network approach examine the linkage between media sentiment and high short interest stocks with emphasis on volatility spillovers. Relevant for our analysis, authors find large drop in connectedness around January 2021, hinting at non-fundamental drivers. The latter two works employ wavelet analysis and more closely examine the GameStop retail trading saga. Umar et al. (2021b) [67], using Twitter and news media data in addition to short-sale volume and options data to proxy retail investor sentiment, document the significant role of sentiment on GameStop returns. The social media and news data however, only consists of publication counts, whereas instead of proxies we mine sentiment directly from the Reddit posts from where the trading frenzy originated. Last, in the paper series, the wavelets framework is utilised to model comovements between the heavily shorted stocks with particular emphasis on GameStop. While Umar et al. (2021c) [68] document strong comovement between GameStop and high short interest indices, more detailed analysis with respect to financial contagion, in particular spillovers to particular assets, is not offered. Similarly, Long, Lucey and Yarovaya (2021) [69] using scraped comments from the r/wallstreetbets subreddit find effect of sentiment on GME returns, sampled at 1-min intervals, albeit limited with ambiguous bidirectional causal linkage.

The strand of literature focusing on the role of individual investors within the GameStop saga largely utilised options data to describe the effect of investor sentiment. Fusari, Jarrow and Lamichhane (2020) [70] utilises options data to identify asset market bubbles, using GameStop bubble at the start of 2021 as an example. On a different note, Jones, Reed and Waller (2021) [71] examines the restrictions imposed on trading and increased margin requirements on specific stocks, showing that the effect of trading restrictions, which is shown different from the effect of mentions on Reddit, led to large price falls with not subsequent recovery. The increased activity in the options market that followed is then shown to accommodate large transfer from option buyers to option writers due to larger prices, open interest and larger implied volatility.

The role of options during the GameStop retail trading saga is further highlighted Allen et al. (2021) [72], where it is shown that options were used to circumvent the trading and short-sale constraints. Further, Baltussen (2021) [73] attempts to explain the extreme dynamics in GME stock and increased activity on option markets by institutional investors' need for so-called *gamma hedging*.

2.6 Hypotheses & Contribution

Textual and semantic analysis in the form of assessing retail investor sentiment with a focus on social media is used in the thesis contributing to the literature examining the role of narratives in economics and financial markets overall. [74] The described trading frenzy presumably driven by retail investor interest depicts the power of the narrative epidemic, in the context of speculative bubbles, which has been greatly amplified by social media where investors accelerated the spread and intensity of the narrative and its life-cycle.

3 Methodology

3.1 Data

Textual and sentiment data

Central to the thesis is the social media data manually scraped from Reddit, in particular the relevant financial subreddits mostly involved in the GameStop saga, especially the *r/wallstreetbets* subreddit. The overview and more information regarding subscriber count are given in Table **??** below.[6] Using Reddit's API service, via API wrappers PRAW and PSAW, we access data on the number of posts dating back to January 2020.[7] Considering the principal role of the *r/wallstreetbets* subreddit, we retrieve data on specific posts corresponding to a set of GME related keywords, including relevant discussion megathreads and daily discussion threads posted on each or preceding trading day. This results in 456,055 unique *r/wallstreetbets* posts spanning from January, 2021 until December, 2021. In the specific posts, there are also included so-called *megathreads* which were created to moderate discussion during most active periods, these threads contain a substantial amount of comments directly related to the most active phases of the GameStop saga, or in reaction to them and thus represent a very concentrated pot of sentiment at the peak of the trading frenzy. Beyond these 59 megathreads, we also gather 419 periodical discussion threads which are posted each day before and after the market open, where users discuss contemporaneous sentiment, strategy and *"future moves"*. Using these threads,

[6] to be supplied
[7] supply links to documentations

we track the general discussion and assess the relevancy of the GME stock by attempting to extract up to 1000 comments and searching for GME related keyword.[8]. For a more detailed description of the data, a small sample of the scraped data is included below. These submissions constitute the ten most upvoted and commented *r/wallstreetbets* posts around the time when GME hit its peak, in January 2021.

Let us briefly discuss one of the limitations of our study. While all submissions must have a descriptive title, we are often not able to extract all the posts' content, beyond the title and other identifying information, as posts may be in the form of images, videos, etc. This is reflected in a significantly lower amount of posts containing a body of text. One possible proxy measure of a post's sentiment could be given by mining the submission's comments. Nevertheless, we primarily rely on the explanatory power of the titles, text submissions and comments in the mentioned mega- and daily threads.

[8]i.e. during the most active phases the activity on the subreddit will be predominantly GameStop related which will be reflected in the daily discussion threads

Table 1: Sample Reddit Posts from January 2021

Most Upvoted Posts

title	score	id	subreddit	num_comments	body	created	datetime
GME YOLO update — Jan 22 2021	91199	l2x7he	wallstreetbets	6332		1.611379e+09	2021-01-23 05:14:08
Honorary WSB Autist award goes to Chamath Pali...	82970	l69jz5	wallstreetbets	2696		1.611797e+09	2021-01-28 01:23:12
GME YOLO update — Jan 25 2021	81959	l4xje1	wallstreetbets	5514		1.611638e+09	2021-01-26 05:05:36
Can I get a flair for buying GME at the litera...	58019	l4sg3u	wallstreetbets	4153		1.611624e+09	2021-01-26 01:18:21
GME YOLO update — Jan 13 2021	52148	kwpviw	wallstreetbets	3188		1.610601e+09	2021-01-14 05:04:19
$500 Donation For Every $50 Increase in GME Pr...	46845	l5be2b	wallstreetbets	1132	[deleted]	1.611686e+09	2021-01-26 18:29:54
GME YOLO update — Jan 19 2021	39183	l0t8ng	wallstreetbets	1615		1.611119e+09	2021-01-20 05:05:55
TRUTH about GME effect!	38375	l67c0p	wallstreetbets	1903		1.611791e+09	2021-01-27 23:51:36
The GME Thread, Part 3.14, for January 27, 2021	36694	l6cb1x	wallstreetbets	49327	Reddit's Engineering team has asked us to rota...	1.611804e+09	2021-01-28 03:26:08
I want to thank you guys for saving my best fr...	36277	l667hb	wallstreetbets	1233	Monday afternoon I took my best friend, an Ame...	1.611788e+09	2021-01-27 23:00:39

Most Commented Posts

title	score	id	subreddit	num_comments	body	created	datetime
The GME Afterhours Thread: Part 4.20 on 27 Jan...	27604	l6er79	wallstreetbets	94621	Stop spamming copy pastes you boomers. Instaba...	1.611812e+09	2021-01-28 05:26:35
The GME Thread Part 1 for January 26, 2021	14456	l5c0nr	wallstreetbets	93889	Good luck today. [Here's some WSB stats.](http...	1.611689e+09	2021-01-26 19:16:56
GME Thoughts, YOLOs, Gains, Stonk Updates,...	18418	l4lmrx	wallstreetbets	93756	Thanks all for the quick rise to max comments ...	1.611601e+09	2021-01-25 18:56:35
GMREEEEEEEEEEEE Containment Thread - GME shitp...	14770	l2ljpt	wallstreetbets	93388	Don't be doxxing citron or anyone else. That's...	1.611342e+09	2021-01-22 19:07:41
GME Megathread Part 2	14249	l4syrd	wallstreetbets	86749	Keep all $GME discussion and memes in here. No...	1.611625e+09	2021-01-26 01:41:39
The GME Thread, Part 2.1, for January 27, 2021	14039	l692dj	wallstreetbets	72294	New thread requested due to the pure chaos our...	1.611796e+09	2021-01-28 01:03:50
GME Megathread - Lemon Party 2: Electric Boogaloo	9072	l1xtan	wallstreetbets	51390	No inauguration today so Citron *may* be able ...	1.611262e+09	2021-01-21 20:49:03
The GME Thread, Part 3.14, for January 27, 2021	36694	l6cb1x	wallstreetbets	49327	Reddit's Engineering team has asked us to rota...	1.611804e+09	2021-01-28 03:26:08
GME Thread: The Wreckoning	12034	l0hhqg	wallstreetbets	46765	Post all your GME hopes, prayers, and stupidit...	1.611082e+09	2021-01-19 18:47:16
GME Megathread - Lemon Party (keep your shitpo...	11861	l19k9r	wallstreetbets	34626	[deleted]	1.611180e+09	2021-01-20 21:52:13

Separately from the scraped data, we also employ the news and social sentiment analysis tools from Bloomberg, downloading news sentiment series for GME, AMC, BB and CLOV, in addition to the Twitter sentiment series for GME. The Twitter series will serve mainly two functions, firstly, to validate our scraped dataset and the mined sentiment. And secondly, it offers a distinctly different point of view offered by another social media, relevant mainly in comparison to news sentiment. It should be noted that while these series are readily usable yielding instant insight, we do not have any background access to the functioning of the sentiment tools and algorithms used to extract sentiment, rendering our comparison possibly weaker. Regardless, the news sentiment datasets are captured on a daily frequency and contain news publication counts which are divided based on the *counts* of positive and negative news sentiments. The same methodology is applied to the Twitter series, the summaries of the datasets are given in Table 6 and Table 5 for GME separately, including Twitter sentiment.

Additionally, for comparison, exogenous measures of overall or financial news sentiment measures are considered, such as sentiment surveys from American Association of Individual Investors (AAII)[9] or WorldData.AI Financial news sentiment index.[10].

Financial data

The historical stock quotes for the relevant tickers are retrieved from several sources based on the frequency. Daily OHLC stock quotes and volume data for the six sampled individual tickers (GME, AMC, BB, NOK, CLOV and RKT) starting from January 2020 are obtained from Yahoo Finance. Our sample choice is primarily driven by the results of the textual analysis, specifically, the mining of most mentioned tickers related to the GME queried posts, especially during the beginning of 2021, that is the sampled companies were most mentioned by posters and commenters in GameStop related threads. In addition to the individual tickers, we also obtain two series of short interest indices from Bloomberg, the Barclays' Short Interest Index (BCCSI) and Citi US Short Interest Equity Index (CIEQUSSI).

Higher frequency, intraday OHLC data based on 30-minute intervals for the sampled individual stocks are likewise downloaded from Bloomberg spanning one year, beginning on November 2020. Using Yahoo Finance's options screener, we manually gather data on GameStop option chains at each of their respective expiries before the market open. Due to manual faults in data collection some data might be missing. Nevertheless, we include the options data for the purposes of validation purposes and cross-checking with other studies which have employed options data to explain the GameStop saga. Furthermore,

[9] `https://www.aaii.com/sentimentsurvey/sent_results`
[10] `https://worlddata.ai/`

Table 2: List of companies

Name	Ticker	CIK	SIC
GameStop	NYSE:GME	1326380	5734 - Retail-Computer & Computer Software Stores
AMC Entertainment	NYSE:AMC	1411579	7830 - Services-Motion Picture Theaters
Blackberry	NYSE:BB	1070235	7372 - Services-Prepackaged Software
Rocket Companies	NYSE:RKT	1805284	6162 - Mortgage Bankers & Loan Correspondents
Clover Health Investments	NASDAQ:CLOV	1801170	6324 - Hospital & Medical Service Plans
Nokia	NYSE:NOK	924613	3663 - Radio & Tv Broadcasting & Communications Equipment

we also have available the historical volatility and implied volatility series from Bloomberg for GME, AMC and BB, the difference between these two series can particularly offer great insight regarding the sentiment implied by speculative trading. Lastly, accompanying the analysis are the short-sale volume statistics which provide the volume of trades classified as short trade on specific trade dates. The FINRA's daily reported *Short Sale Transaction Files* [11], or indirectly from the streamlined datasets for individual tickers available on Quandl [12]. The overall summary statistics for the daily and intraday 30-min returns and volatility series are given in Table 3 and Table 4, respectively.

Standardisation ...

Table 3: Return Series Data

	GME	AMC	BB	NOK	CLOV	RKT	BCCSI	CIEQUSSI
Daily Data								
N	512	512	512	512	389	351	261	512
Mean	0.0134	0.0102	0.0023	0.0016	-0.0005	0.0006	0.0004	0.0001
Std	0.1302	0.1679	0.0549	0.0346	0.0671	0.0547	0.0275	0.0062
Min	-0.6000	-0.5663	-0.4163	-0.2840	-0.2361	-0.3267	-0.1448	-0.0415
25%	-0.0312	-0.0402	-0.0225	-0.0114	-0.0243	-0.0170	-0.0146	-0.0030
50%	-0.0011	-0.0069	-0.0021	0.0000	-0.0015	-0.0009	0.0000	0.0000
75%	0.0342	0.0370	0.0209	0.0137	0.0129	0.0119	0.0138	0.0031
Max	1.3484	3.0121	0.3266	0.3848	0.8582	0.7119	0.2099	0.0478
Kurtosis	36.7527	202.3506	14.2821	40.4758	70.7780	85.5108	15.7279	14.1397
Skewness	4.1682	11.6602	0.3869	1.2984	5.7956	5.8473	1.2770	0.5376
J-B stat	29713.4008	868012.5544	4272.8479	34395.8250	81255.6575	105867.5163	2651.0264	4200.0410
Intraday Data								
N	3170	3170	3170	3170	3170	3170.		
Mean	0.0018	0.0014	0.0004	0.0001	0.0001	0.0000		
Std	0.0469	0.0393	0.0189	0.0098	0.0192	0.0122		
Min	-0.5000	-0.3942	-0.2543	-0.2036	-0.1139	-0.2405		
25%	-0.0082	-0.0089	-0.0053	-0.0024	-0.0064	-0.0035		
50%	-0.0002	-0.0005	-0.0003	0.0000	-0.0007	-0.0003		
75%	0.0078	0.0078	0.0047	0.0020	0.0055	0.0030		
Max	1.1259	1.3200	0.3254	0.2839	0.4841	0.1955		
Kurtosis	182.2834	419.9672	79.4065	304.1448	136.7407	95.2830		
Skewness	7.9663	12.8059	3.6964	6.5254	5.8380	-0.4265		
J-B stat	4.41e+06	2.33e+07	8.37e+05	1.22e+07	2.48e+06	1.20e+06		

[11] `https://www.finra.org/finra-data/browse-catalog/short-sale-volume-data/monthly-short-sale-volume-files`

[12] `https://data.nasdaq.com/data/FINRA-financial-industry-regulatory-authority`

Table 4: Volatility Series

	GME		AMC		BB	
	Hist Vol	Implied Vol	Hist Vol	Implied Vol	Hist Vol	Implied Vol
N	252	252	252	252	252	251
Mean	158.6027	160.7568	158.6166	175.4694	83.9137	95.6295
Std	174.6587	90.7542	173.4695	76.1244	67.4837	48.8617
Min	27.4771	68.2247	26.6585	94.1729	22.3755	52.6335
25%	58.0469	102.6217	77.4897	130.2942	46.8784	67.8324
50%	88.8131	130.1494	114.2138	157.9829	66.2621	80.2687
75%	167.6699	176.9667	163.7235	189.1723	88.7712	106.7923
Max	999.0340	538.3421	1016.6746	779.4945	403.9863	454.5569
Kurtosis	9.0113	3.4489	14.2485	18.4284	11.9106	16.6268
Skewness	2.8161	1.8656	3.6691	3.3703	3.2824	3.5035
J-B stat	1144.0467	262.7743	2600.1436	3888.9561	1872.6817	3277.0221

Table 5: GME Sentiment Series

	News				Twitter		
	Count	Positive	Negative	Net	Count	Positive	Negative
N	521	519	521	522	521	508	514
Mean	143.6564	6.8112	-8.5470	-1.7586	2312.2745	146.0807	-183.5661
Std	423.4256	22.1425	34.3938	29.4166	6814.4936	420.1067	568.4365
Min	1.0000	0.0000	-438.0000	-291.0000	17.0000	1.0000	-5960.0000
25%	13.0000	0.0000	-3.0000	-1.0000	104.0000	7.0000	-116.7500
50%	36.0000	1.0000	-1.0000	0.0000	609.0000	39.0000	-47.0000
75%	106.0000	4.0000	0.0000	1.0000	1751.0000	103.0000	-10.0000
Max	4587.0000	235.0000	0.0000	223.0000	74211.0000	4233.0000	-1.0000
Kurtosis	61.8958	61.4673	82.9939	59.4741	59.2965	49.0166	52.3648
Skewness	7.2652	7.1507	-8.3633	-3.9709	7.0511	6.4555	-6.7508

Table 6: Sentiment Series

	AMC			BB			CLOV		
	News Count	PosSent	NegSent	News Count	PosSent	NegSent	News Count	PosSent	NegSent
N	521	521	521	511	517	521	338	255	244
Mean	138.8522	5.8714	-6.1574	27.9217	1.3926	-1.3800	21.4172	1.4549	-2.1148
Std	246.5869	17.6954	16.6840	43.0047	4.7799	4.4051	44.7808	8.5365	4.5744
Min	3.0000	0.0000	-200.0000	1.0000	0.0000	-44.0000	1.0000	0.0000	-40.0000
25%	39.0000	0.0000	-5.0000	7.0000	0.0000	-1.0000	4.0000	0.0000	-2.0000
50%	74.0000	1.0000	-2.0000	13.0000	0.0000	0.0000	10.5000	0.0000	0.0000
75%	140.0000	4.0000	0.0000	30.0000	1.0000	0.0000	25.0000	0.0000	0.0000
Max	2631.0000	261.0000	0.0000	354.0000	48.0000	0.0000	545.0000	118.0000	0.0000
Kurtosis	41.1196	96.8206	75.8717	19.1666	44.9040	41.9799	88.9844	146.0938	28.3862
Skewness	5.7473	8.3008	-7.6792	3.9191	6.2268	-6.0105	8.4732	11.4648	-4.5908

4 Results

4.1 Text Analysis

Firstly, we begin with a purely semantic analysis. Using the whole sample of more than 438 thousand r/wallstreetbets unique submissions, word tokens, Porter's stems and lemmas are extracted. The overwhelming positive sentiment is evident at first sight as shown by the Fig. [13], depicting the most frequent word tokens and bigrams. The prominency of words *rocket, buy, hold, moon* or *let* tokens indicates a large bullish sentiment. Broader context is presented by the bigram model extracting phrases such as *rocket rocket, let us, moon rocket, us go* or *hold line*[14]. Fig. **??** shows the activity on the subreddit throughout the year corresponding to most volatile periods of GME.

Moreover, highly frequent *short-squeeze, shares* or *short interest* terms manifest the relevancy of the short-selling phenomena in the trading frenzy, it arguably being one of its triggering factors. The relevance of the selected sample of companies, as shown by Fig. **??**, demonstrates the large association users made between GameStop and other tickers. The key to identifying ticker mentions is based on a simple convention that ticker must be preceded by a $ sign and contain up to six letters.

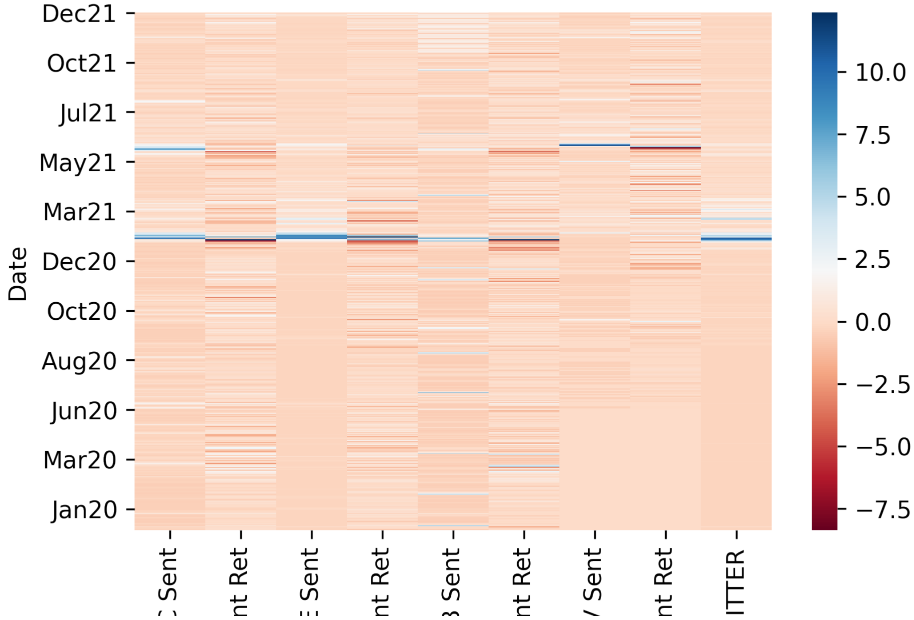
Sentiment Analysis

Moving beyond a simple textual analysis, we assign sentiment scores to the scraped submissions. Owing to the specific lexicon predominantly used in the r/wallstreetbets subreddit, we believe that slight modifications of the lexicon are necessary to better reflect bearish and bullish signals.[15] Therefore, we construct two sentiment scores, before and after applying the lexicon adjustments, the overview of which is given in Appendix **??**. Our analysis primarily employs the latter.

Fig. **??** further confirm the two distinct periods in which sentiment surrounding

[13] number
[14] translation of emojis may distort the most used phrases
[15] overview of some of the changes will be referred here, in appendix

Figure 1: Heatmaps



GameStop peaked on the subreddit. It is shown that positive sentiment dominated throughout the year, regardless of the performance and volatility of GME. Interestingly, aggregating the scores intraday, we detect large deviations and periods of strong negative sentiment, which almost disappear in aggregations of frequencies larger than 6 hours.[16] This may suggest a fickle and reactive behaviour on the subreddit as negative mood is quickly suppressed and drown out by the trend of overwhelming positive sentiment.[17].

Comparing the news and Reddit sentiments, we observe very different patterns however, note the large discrepancy in the respective datasets' sizes. While the former shows significant peaks and troughs, Reddit users' sentiment seems to be unwaveringly positive throughout the period. On the other hand, comparing the two social media sentiment measures from Twitter and Reddit, we observe an even more stark contrast. The Twitter sentiment displays more or less consistently opposite signals compared to the Reddit sentiment, being largely negative especially during the most active periods.

Nevertheless, all sentiment datasets pick on the increased activity and heightened volatility since the beginning of 2021 corresponding to the GameStop trading frenzy. Likewise, the news sentiment surrounding the other relevant tickers displays similar patterns, as shown by Fig. **??**.

[16]either figure included here or comparison in appendix
[17]echo chamber, repetition?

Figure 2: Word & Bigram Frequency

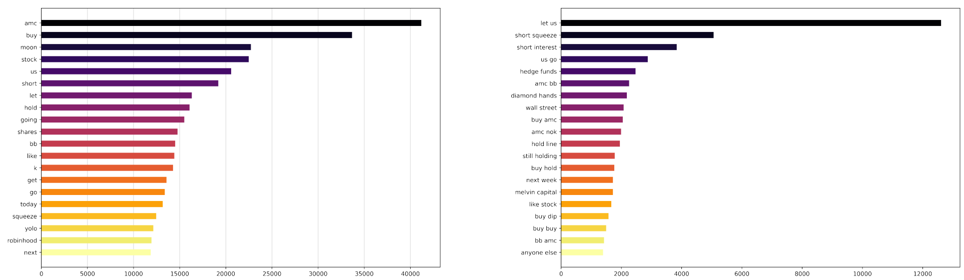
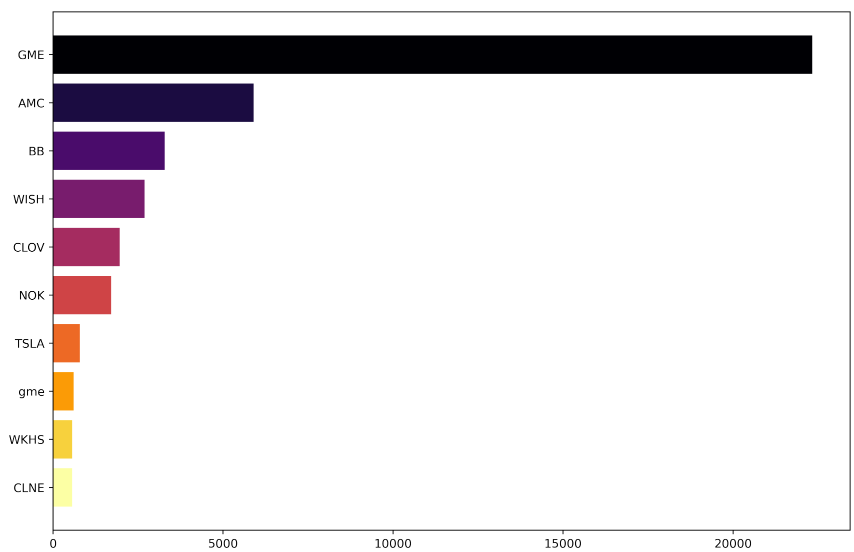


Figure 3: Most Mentioned Tickers



4.2 Wavelet Coherence

Let us now discuss the core findings of our analysis. Applying the wavelet framework to analyse the effects of Reddit investors' sentiment during the GameStop retail trading frenzy, we discover interesting dynamics. Figure [18] shows wavelet coherence plots between the scraped sentiment and intraday 30-min returns, historical and implied volatility and lastly, volume.[19]. The colour scale on the right-hand side corresponds to the level of wavelet coherence with warmer (colder) colours signifying higher (lower) coherence. The arrows represent the relative phase relationship between the two series which can be indicative of causality. The arrow notation follows Torrence and Campo (1998) [65]. In-phase (anti-phase) behaviour is denoted as $\uparrow$ ($\downarrow$). In-phase and anti-phase relationships can be alternatively described as positive and negative correlations. The phase difference is then indicated by the horizontal plane of the arrows' notation where a right-pointing $\rightarrow$ ($\leftarrow$) arrows suggest that the first (second) series leads the second (first) by $90°$. For illustration, in Panel A $\nearrow$ implies that sentiment and GME returns are in-phase and the former leads the latter by $45°$ in other words, sentiment positively influences GME returns. The horizontal axis corresponds to the time dimension while the vertical axis represents the frequency dimension, respectively inverse period. The scales (periods) range from 2 up to 128 intervals which depend on the sampling frequency, i.e. 1 to 64 hours[20] or 2 to 128 days. White contours designate time-scale areas where wavelet coherence is statistically significant at a 5% level given by the Monte Carlo simulations procedure described above. Finally, the cone of influence is given by the cross-hatched region wherein the edge effects can be effectively ignored.

Panel A shows number of short-lived areas of significant coherence at higher frequencies between the sentiment and GME return series, which can however, occur by chance. More telling are the large areas of coherence between the 8-128 scales the timing of which corresponds to the most active periods in the subreddit. The two substantial periods of significant coherence occurring in Jan-Mar and May-Jun periods show two series being largely in phase with sentiment leading, suggesting that in scales from 8 hours up to 64 hours, corresponding to approx. 1 to 8 trading days, retail investors sentiment positively influenced GME returns. Further, beginning August there are also larger however shorter-lived areas of significant coherence at high frequencies, notably in the 8-32 frequency bands. Although, we observe a phase switch as in these time-scale regions both series are interestingly anti-phase with no distinct phase difference. That is, increased activity larger sentiment had no effect on the stock's returns and moreover, larger sen-

[18] numref

[19] figure note that some series are sampled at different frequencies

[20] remember 30-min intervals

timent was simply met with lower returns which is in contrast to previous periods of heightened activity on the subreddit.

Contrary to our expectation that sentiment is better predictor of volatility rather than series, we observe fewer and smaller time-scale areas of significant wavelet coherence, as shown by Panel B. Nonetheless, during the most active periods we can likewise identify larger areas of significant dependence on the 16-128 scales corresponding to 8-64 hour frequencies. Sentiment and intraday volatility is largely in-phase with a small bias towards the former leading. Although, there are some periods where the phase-difference switches as volatility leads sentiment , notably in June and towards the end of the year.

We also employ daily data on implied volatility and traded volume, presented in Panels C and D, respectively. Given the relatively short time window for which we have available data (1 year) we are mostly interested in higher frequencies. Panel C shows large time-scale regions of coherence on 2-8 scales (days), which somewhat coincide with the large volatility spikes of GME. Furthermore, the phase difference indicates that sentiment is positively influenced by volatility, sustaining the argument that greater volatility simply spurs more discussion and feeds the growing sentiment on the subreddit. At lower frequencies (¿16 days) we observe consistent in-phase coherence which however, disappears after August despite the two series exhibiting dependence at higher frequencies in that time period.

Finally, we also assess the dependence between sentiment and traded volume. The activity on the subreddit and the market are found coherent, even at higher frequencies at times in the 2-8 frequency bands, with the sentiment being positively influenced by the traded volume. The dependence at lower frequencies beyond 16 days shows that at greater scales the subreddit activity is merely given by the activity on the market. This indicates that retail investors' discussion may present an echo-chamber wherein fundamental information presumably, the primary driver of increased activity, is mulled over, creating substantial noise. Essentially, the revealed dependence suggests that heightened trading activity leads to retail participants discussing more and further fuelling the sentiment on the subreddit.

News Sentiment

Provided that the news dataset compared to the Reddit scraped sentiment extends further to December 2019, it additionally supplements unique findings. Similarly, we observe large time-scale regions across all frequencies with significant coherence, especially from September 2020 onwards. Note that until the end of 2020 the phase relationship saw primarily news lead GME returns. Following that the phase difference changes direction as GME returns positively influenced news sentiment in the first three months of 2021

Figure 4: PANEL A

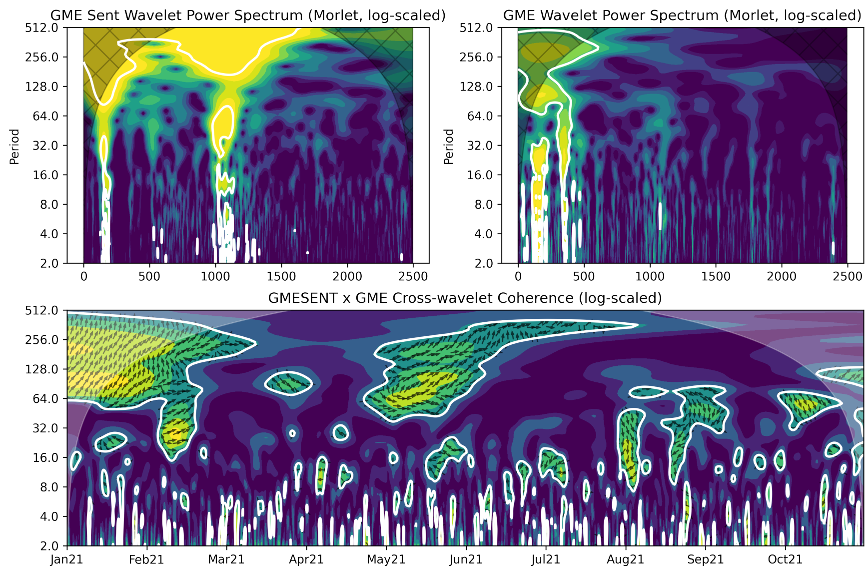


Figure 5: Panel B

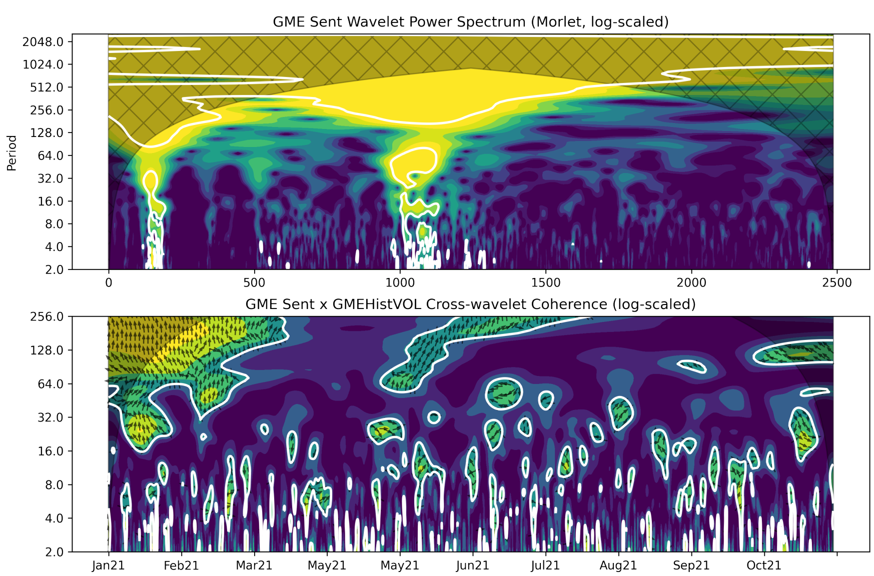


Figure 6: Panel C

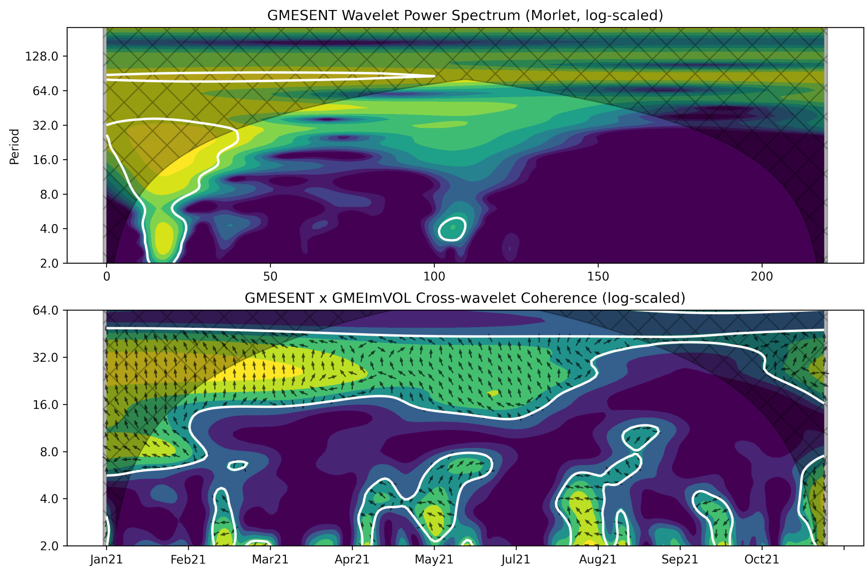


Figure 7: Panel D

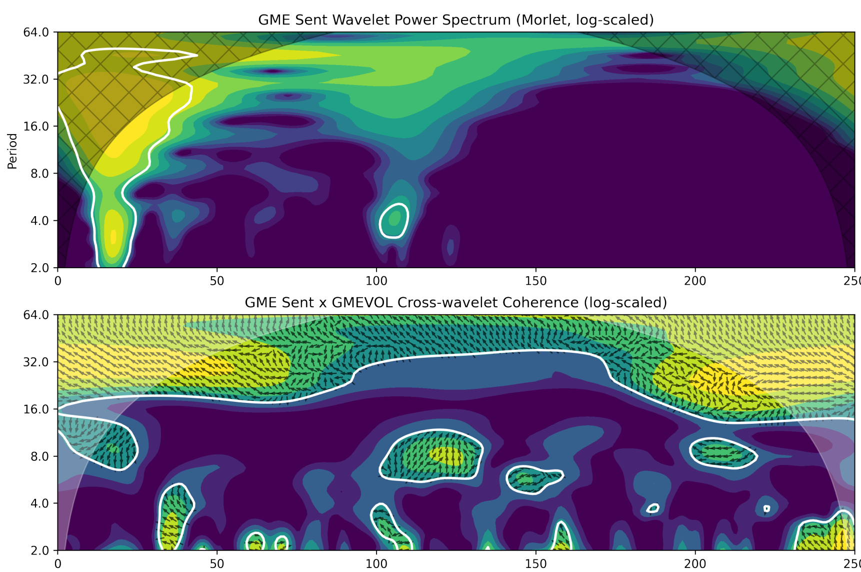


Figure 8: Panel D

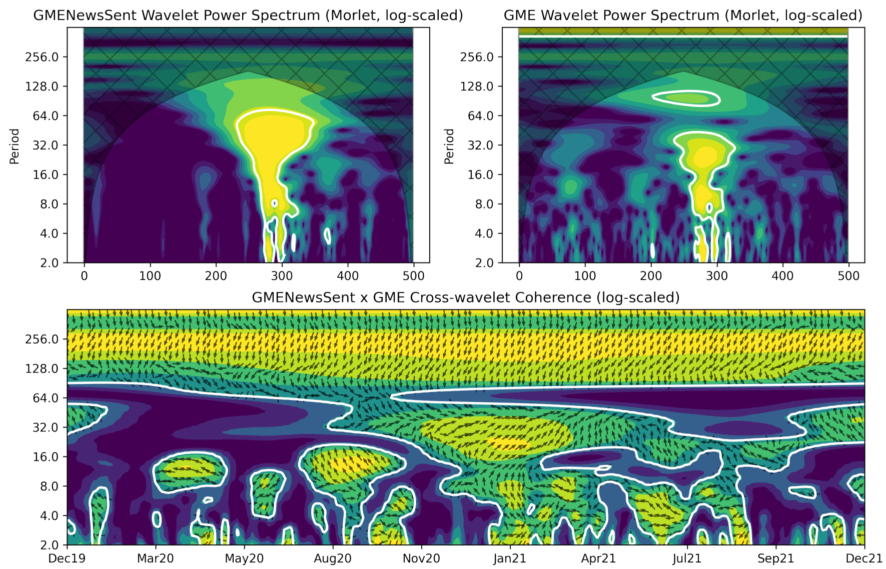
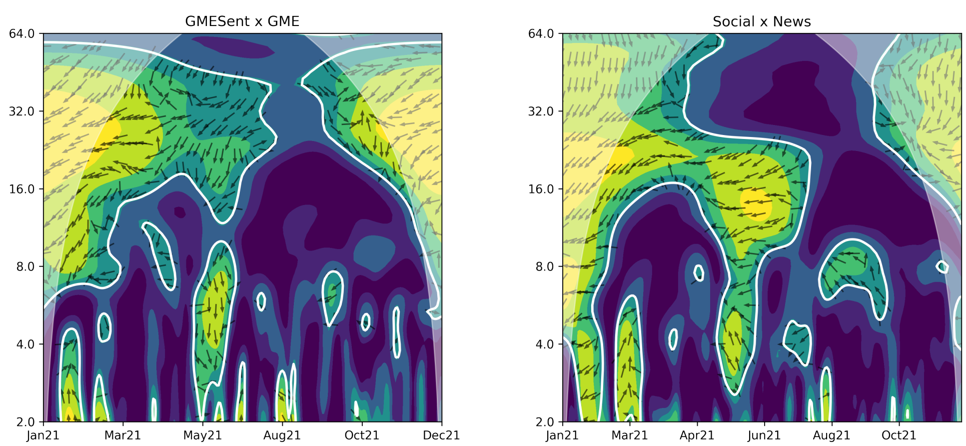


Figure 9: Daily Sentiment



when the GameStop saga peaked. Later that year, we can still identify sizeable regions of significant dependence showing that GME returns mainly reacted to news sentiment as phase difference reverted. The structural break in the possible causality may suggest that while otherwise, the stock responded to presumably fundamental information, at the beginning of the year, the presence of a different driver as retail investors took the reins.

Furthermore, allowing a more direct comparison, we re-run the wavelet coherence using the daily aggregated Reddit sentiment dataset. The time localisation of the large regions of significant coherence does not change. Significant dependence at higher frequencies at 2-16 scales (days) is also found during the peak activity periods. The phase arrows, however, indicate that phase difference is in favour of GME returns, suggesting that they lead the sentiment on the subreddit. Meaning that the retail investors' sentiment was largely in reaction to the market development, contrary to what we have observed above. Although note the difference between the scales in the figures, for instance, 8-32 scales (days) in Fig. 9 correspond to approx. 384-1536 scales in Fig. 4. Moreover, in longer investment horizons, it is shown that sentiment and GME returns exhibit an anti-phase relationship. One possible explanation for the clashing results given by different sampling frequencies is that while retail investors' sentiment could drive returns in very short investment horizons up to a few days, on a longer time frame their impact is limited.[21]

Lastly, Panel X plots the wavelet coherence between the two sentiments sampled on daily frequencies. We identify significant dependence at higher frequencies up to 16 days occurring during the peak periods, namely around the Jan-Mar and May-June periods. Interestingly, while the two sentiment series are in-phase in some high-frequency regions, they are anti-phase throughout most of the sample period. The phase difference indicates that it is predominantly news sentiment determining the sentiment on the subreddit in longer time horizons. Thus, highlighting the interpretation role of social media and the echo chamber effect.

4.3 Comovements

Finally, we also employ wavelets framework to analyse the contagion of the GameStop saga relevant to the selected sample of stocks discussed earlier. In addition to these most mentioned tickers surrounding the GME discussion on the subreddit, market-wide and short interest indices are used to assess broader contagion. The results are reported in the respective Panels A-Z in Fig. X **??**. AMC and GME exhibit the largest comovement by far from the sampled stocks as shown by the predominantly warm coloured time-scale regions in Panel X, indicating large and consistent coherence. Both stocks show very

[21] IN REVISION

high degree of dependence even at higher frequencies, which is interestingly even more intense following June. As expected considering the results of the textual analysis both are in-phase. However, there is no distinct phase difference between them, suggesting that the possible spillovers went both ways, feeding off each other.

The other stocks do not display even nearly the same level of comovement with GME as AMC. Nevertheless, there are large significant time-scale regions where comovement occurs even at higher frequencies around 4-16 scales (2-8hours). These time periods primarily coincide with the most active periods in the GME saga as well as the volatility spikes of the respective stocks. All series predominantly exhibit in-phase relationship. Moreover, there is no unequivocal evidence as to whether spillovers came from GME or the other way around, as indicated by the phase arrows. Therefore, apart from AMC, the contagion is largely limited to the periods of higher activity on the subreddit and during which these stocks experienced large price and volatility spikes. While the comovement at higher frequencies is relatively confined, the results suggest that retail investors activity could have led volatility and return spillovers.

Let us now turn to a broader context examining the market-wide and short interest indices comovements. GME and the short interest index do not exhibit larger significant comovement until the end of 2020. Already starting from October 2020 we can see significant comovement which expands to a larger frequency band about 2-16 days. Interestingly, during this period, at lower frequencies the short index leads GME while at higher frequencies phase difference is opposite. Despite significant short-interest in GME we are not able to identify strong and consistent in- or anti-phase relationship although, for a brief period in July, there is a strong anti-phase comovement. The differing phase difference and lack of phase relationship might lay evidence that GME returns were hugely driven, albeit for a short period, by retail investors' activity leading to large pricing deviations from the fundamentals.

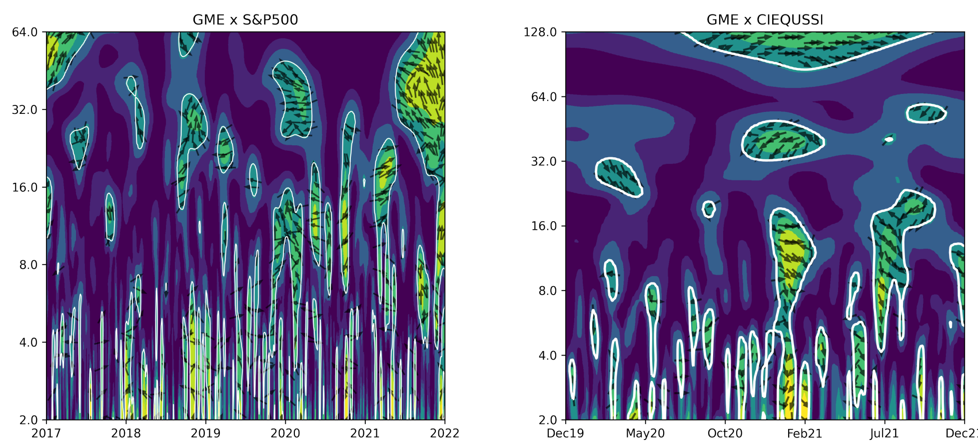
Historically, there is understandably significant comovement at even higher frequencies although relatively short-lived. The increase in comovement also corresponds to larger market declines when market complexity is reduced as multi-correlation between all stocks unilaterally increases as the common drivers arise such as macroeconomic conditions. This is in fact documented by xxx [22]. Besides the large coherent areas in 2020, starting from 2021 we observe a large white contoured time-scale region signalling in-phase dependence wherein GME leads. Nonetheless, given the historical pattern of comovement and sparsity of large significant time-scale region we cannot conclude that GME saga spilled over to the broad market.

In summary, we identify large significant time-scale regions where GME and other

[22] refer to entropy study

relevant stocks exhibit strong in-phase comovement and these periods coincide with the peaking activity on the subreddit. Furthermore, unsurprisingly during these periods each respective stock saw large volatility and their prices at the peak. AMC in particular displays extremely strong and consistent comovement throughout the whole sample period across all frequencies, which is a unique finding in our sample. Although the phase differences do not bear conclusive evidence regarding the direction of the spillovers, using the short interest index and broad market index, it is shown that the comovements at the times they occur are unique and not a result of broad market sentiment. Rather, we suggest that these spillovers leading to large pricing deviations and volatility spikes are uniquely determined by the growing sentiment of the retail investors on the subreddit.

Figure 10:
Note the difference in sample period. The left-hand plot uses daily frequency data over the last 5 years, while the right hand plot plots comovement between GME and index tracking basket of shorted stocks sampled at daily frequency from December 2019 until December 2021.



Following our supposition regarding sentiment being a better predictor of volatility rather than returns, we assess volatility spillovers in addition to returns, the figures are included in Appendix. The results are mostly similar, we do not observe any substantial departure from the above findings. The coherence plots show abundant short-lived time-scale regions of significant comovement at higher frequencies which however, could simply occur by chance or result from broad market sentiment. The larger and longer time-scale regions at higher wavelengths correspond to the same regions as in return series. The larger areas stretching over all relevant frequencies show that higher GME volatility accorded with higher volatility of other stocks however, likewise there is not conclusive evidence that spikes in GME volatility spilled over and directly caused spikes in the other discussed stocks.

Figure 11: Volatility

Description: Note the difference in sample period. The left-hand plot uses daily frequency data over the last 5 years, while the right hand plot plots comovement between GME and index tracking basket of shorted stocks

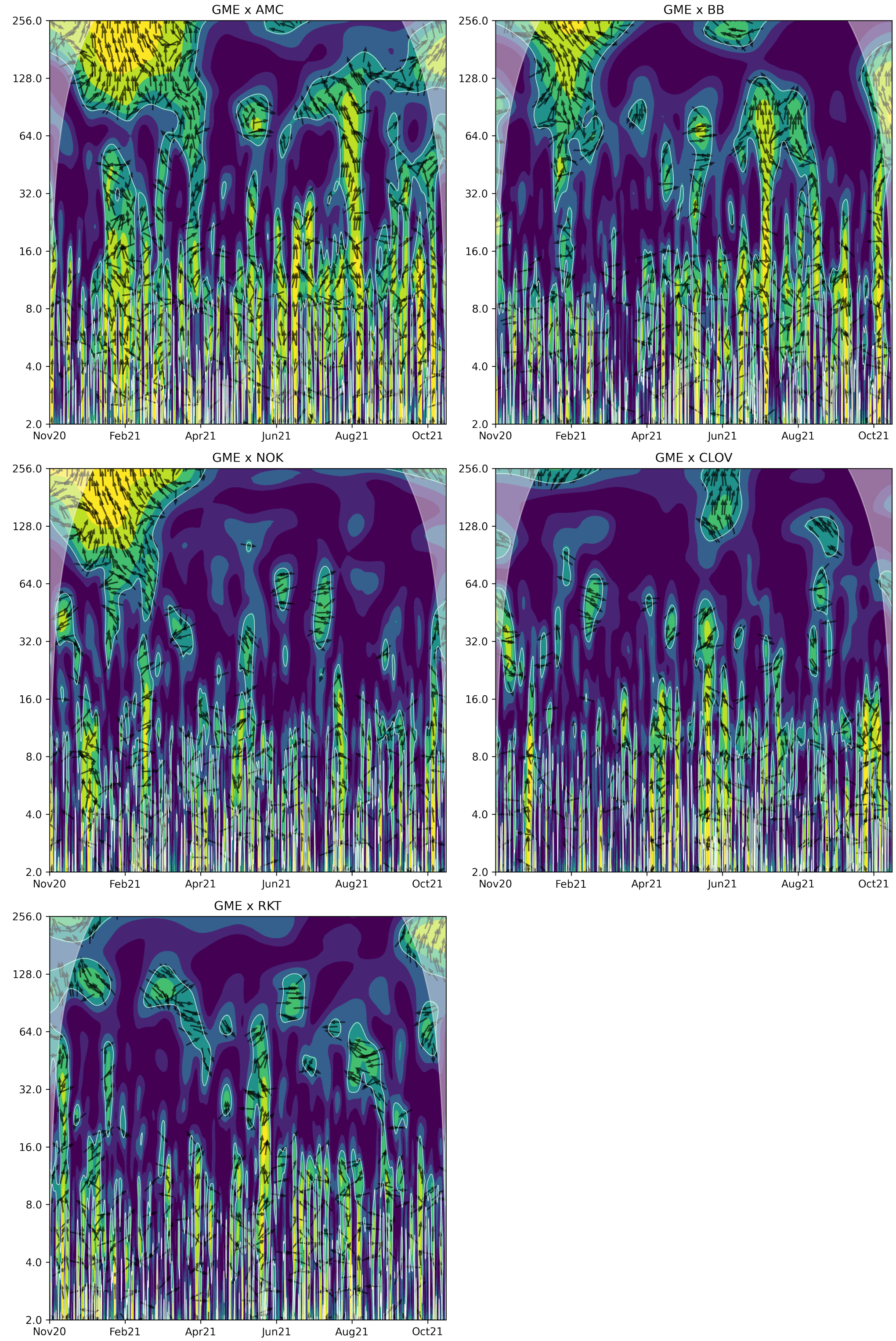
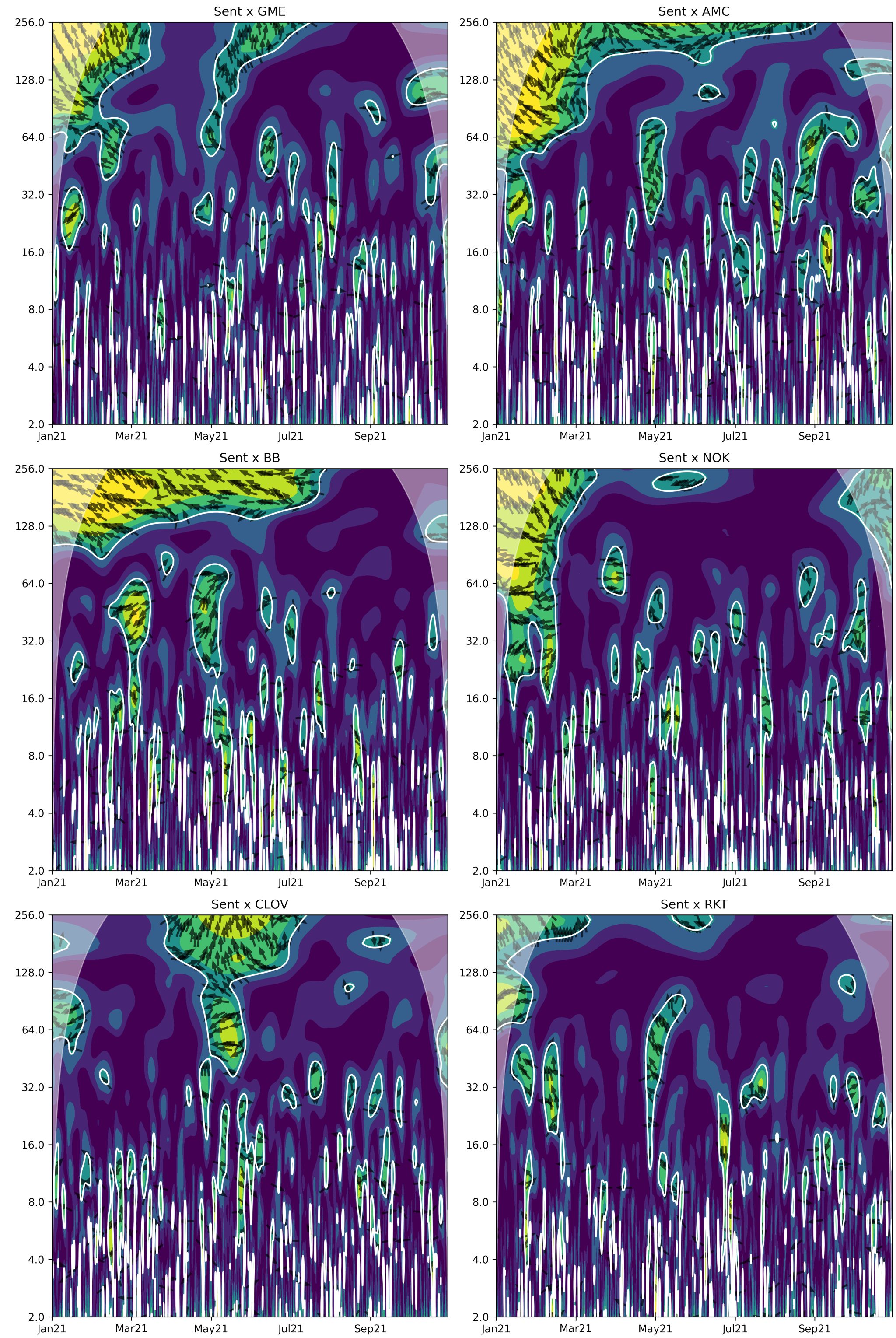


Figure 12: Sentiment x Volatility



5 Discussion

Conclusion

References

[1] E. F. Fama, "Efficient capital markets: A review of theory and empirical work," *The Journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970.

[2] G. Yen and C.-f. Lee, "Efficient market hypothesis (emh): Past, present and future," *Review of Pacific Basin Financial Markets and Policies*, vol. 11, no. 02, pp. 305–329, 2008.

[3] R. J. Shiller, "From efficient markets theory to behavioral finance," *The Journal of Economic Perspectives*, vol. 17, no. 1, pp. 83–104, 2003.

[4] W. F. M. D. Bondt and R. Thaler, "Does the stock market overreact?," *The Journal of Finance*, vol. 40, no. 3, pp. 793–805, 1985.

[5] M. S. Rozeff and W. R. Kinney, "Capital market seasonality: The case of stock returns," *Journal of Financial Economics*, vol. 3, no. 4, pp. 379–402, 1976.

[6] R. Roll, "Vas ist das? the turn-of-the-year effect and the return premia of small firms," *The Journal of Portfolio Management*, vol. 9, no. 2, pp. 18–28, 1983.

[7] M. R. Reinganum, "The anomalous stock market behavior of small firms in january: Empirical tests for tax-loss selling effects," *Journal of Financial Economics*, vol. 12, no. 1, pp. 89–104, 1983.

[8] J. R. Ritter, "The buying and selling behavior of individual investors at the turn of the year," *The Journal of Finance*, vol. 43, no. 3, pp. 701–717, 1988.

[9] A. S. Kyle, "Continuous auctions and insider trading," *Econometrica*, vol. 53, no. 6, pp. 1315–1335, 1985.

[10] F. Black, "Noise," *The Journal of Finance*, vol. 41, no. 3, pp. 529–543, 1986.

[11] A. Shleifer and L. H. Summers, "The noise trader approach to finance," *The Journal of Economic Perspectives*, vol. 4, no. 2, pp. 19–33, 1990.

[12] J. B. D. Long, A. Shleifer, L. H. Summers, and R. J. Waldmann, "The size and incidence of the losses from noise trading," *The Journal of Finance*, vol. 44, no. 3, pp. 681–696, 1989.

[13] J. B. de Long, A. Shleifer, L. H. Summers, and R. J. Waldmann, "Positive feedback investment strategies and destabilizing rational speculation," *The Journal of Finance*, vol. 45, no. 2, pp. 379–395, 1990.

[14] J. B. D. Long, A. Shleifer, L. H. Summers, and R. J. Waldmann, "Noise trader risk in financial markets," *Journal of Political Economy*, vol. 98, no. 4, pp. 703–738, 1990.

[15] R. Mehra and E. C. Prescott, "The equity premium: A puzzle," *Journal of Monetary Economics*, vol. 15, no. 2, pp. 145–161, 1985.

[16] C. M. C. Lee, A. Shleifer, and R. H. Thaler, "Investor sentiment and the closed-end fund puzzle," *The Journal of Finance*, vol. 46, no. 1, pp. 75–109, 1991.

[17] G. W. Brown, "Volatility, sentiment, and noise traders," *Financial Analysts Journal*, vol. 55, no. 2, pp. 82–90, 1999.

[18] J. N. Bodurtha, D.-S. Kim, and C. M. C. Lee, "Closed-end country funds and u.s. market sentiment," *The Review of Financial Studies*, vol. 8, no. 3, pp. 879–918, 1995.

[19] J. Pontiff, "Excess volatility and closed-end funds," *The American Economic Review*, vol. 87, no. 1, pp. 155–169, 1997.

[20] R. J. Barkham and C. W. R. Ward, "Investor sentiment and noise traders: Discount to net asset value in listed property companies in the u.k.," *The Journal of Real Estate Research*, vol. 18, no. 2, pp. 291–312, 1999.

[21] B. Swaminathan, "Time-varying expected small firm returns and closed-end fund discounts," *The Review of Financial Studies*, vol. 9, no. 3, pp. 845–887, 1996.

[22] S. Kothari and J. Shanken, "Book-to-market, dividend yield, and expected market returns: A time-series analysis," *Journal of Financial Economics*, vol. 44, no. 2, pp. 169–203, 1997.

[23] R. Neal and S. M. Wheatley, "Do measures of investor sentiment predict returns?," *The Journal of Financial and Quantitative Analysis*, vol. 33, no. 4, pp. 523–547, 1998.

[24] M. Baker and J. Wurgler, "The equity share in new issues and aggregate stock returns," *The Journal of Finance*, vol. 55, no. 5, pp. 2219–2257, 2000.

[25] M. Baker and J. Wurgler, "Investor sentiment and the cross-section of stock returns," *The Journal of Finance*, vol. 61, no. 4, pp. 1645–1680, 2006.

[26] M. Baker and J. Wurgler, "Investor sentiment in the stock market," *The Journal of Economic Perspectives*, vol. 21, no. 2, pp. 129–151, 2007.

[27] C. M. Lee, "Earnings news and small traders: An intraday analysis," *Journal of Accounting and Economics*, vol. 15, no. 2, pp. 265–302, 1992.

[28] T. Odean, "Do investors trade too much?," *The American Economic Review*, vol. 89, no. 5, pp. 1279–1298, 1999.

[29] B. M. Barber and T. Odean, "Trading is hazardous to your wealth: The common stock investment performance of individual investors," *The Journal of Finance*, vol. 55, no. 2, pp. 773–806, 2000.

[30] B. M. Barber and T. Odean, "Boys will be boys: Gender, overconfidence, and common stock investment," *The Quarterly Journal of Economics*, vol. 116, no. 1, pp. 261–292, 2001.

[31] C.-C. Chang, P.-F. Hsieh, and Y.-H. Wang, "Sophistication, sentiment, and misreaction," *The Journal of Financial and Quantitative Analysis*, vol. 50, no. 4, pp. 903–928, 2015.

[32] B. M. Barber and T. Odean, "Online investors: Do the slow die first?," *The Review of Financial Studies*, vol. 15, no. 2, pp. 455–487, 2002.

[33] A. Jackson, "The aggregate behaviour of individual investors," *SSRN Electronic Journal*, 2003.

[34] L. Feng and M. S. Seasholes, "Correlated trading and location," *The Journal of Finance*, vol. 59, no. 5, pp. 2117–2144, 2004.

[35] N. Barberis, A. Shleifer, and J. Wurgler, "Comovement," *Journal of Financial Economics*, vol. 75, no. 2, pp. 283–317, 2005.

[36] B. M. Barber and T. Odean, "All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors," *The Review of Financial Studies*, vol. 21, no. 2, pp. 785–818, 2008.

[37] A. Kumar and C. M. C. Lee, "Retail investor sentiment and return comovements," *The Journal of Finance*, vol. 61, no. 5, pp. 2451–2486, 2006.

[38] B. M. Barber, T. Odean, and N. Zhu, "Do retail trades move markets?," *The Review of Financial Studies*, vol. 22, no. 1, pp. 151–186, 2009.

[39] R. Kaniel, G. Saar, and S. Titman, "Individual investor trading and stock returns," *The Journal of Finance*, vol. 63, no. 1, pp. 273–310, 2008.

[40] T. FOUCAULT, D. SRAER, and D. J. THESMAR, "Individual investors and volatility," *The Journal of Finance*, vol. 66, no. 4, pp. 1369–1406, 2011.

[41] S. C. Andrade, C. Chang, and M. S. Seasholes, "Trading imbalances, predictable reversals, and cross-stock price pressure," *Journal of Financial Economics*, vol. 88, no. 2, pp. 406–423, 2008.

[42] M. W. Brandt, A. Brav, J. R. Graham, and A. Kumar, "The idiosyncratic volatility puzzle: Time trend or speculative episodes?," *The Review of Financial Studies*, vol. 23, no. 2, pp. 863–899, 2010.

[43] M. J. Kamstra, L. A. Kramer, and M. D. Levi, "Winter blues: A sad stock market cycle," *The American Economic Review*, vol. 93, no. 1, pp. 324–343, 2003.

[44] A. Edmans, D. García, and Ø. Norli, "Sports sentiment and stock returns," *The Journal of Finance*, vol. 62, no. 4, pp. 1967–1998, 2007.

[45] W. N. Goetzmann, D. Kim, A. Kumar, and Q. Wang, "Weather-induced mood, institutional investors, and stock returns," *The Review of Financial Studies*, vol. 28, no. 1, pp. 73–111, 2015.

[46] A. Edmans, A. Fernandez-Perez, A. Garel, and I. Indriawan, "Music sentiment and stock returns around the world," *Journal of Financial Economics*, 2021.

[47] D. M. Cutler, J. M. Poterba, and L. H. Summers, "What moves stock prices?," *The Journal of Portfolio Management*, vol. 15, no. 3, pp. 4–12, 1989.

[48] G. Gidofalvi and C. Elkan, "Using news articles to predict stock price movements," *Department of Computer Science and Engineering, University of California, San Diego*, 2001.

[49] G. Fung, J. Yu, and H. Lu, "The predicting power of textual information on financial markets.," *IEEE Intelligent Informatics Bulletin*, vol. 5, pp. 1–10, 01 2005.

[50] P. C. Tetlock, "All the news that's fit to reprint: Do investors react to stale information?," *The Review of Financial Studies*, vol. 24, no. 5, pp. 1481–1512, 2011.

[51] R. P. Schumaker and H. Chen, "Textual analysis of stock market prediction using breaking financial news: The azfin text system," *ACM Trans. Inf. Syst.*, vol. 27, mar 2009.

[52] Z. DA, J. ENGELBERG, and P. GAO, "In search of attention," *The Journal of Finance*, vol. 66, no. 5, pp. 1461–1499, 2011.

[53] Z. Da, J. Engelberg, and P. Gao, "The Sum of All FEARS Investor Sentiment and Asset Prices," *The Review of Financial Studies*, vol. 28, pp. 1–32, 10 2014.

[54] T. Preis, H. S. Moat, and H. E. Stanley, "Quantifying trading behavior in financial markets using google trends," *Scientific Reports*, vol. 3, no. 1, p. 1684, 2013.

[55] H. S. Moat, C. Curme, A. Avakian, D. Y. Kenett, H. E. Stanley, and T. Preis, "Quantifying wikipedia usage patterns before stock market moves," *Scientific Reports*, vol. 3, no. 1, p. 1801, 2013.

[56] W. Antweiler and M. Z. Frank, "Is all that talk just noise? the information content of internet stock message boards," *The Journal of Finance*, vol. 59, no. 3, pp. 1259–1294, 2004.

[57] H. Chen, P. De, Y. J. Hu, and B.-H. Hwang, "Wisdom of crowds: The value of stock opinions transmitted through social media," *The Review of Financial Studies*, vol. 27, no. 5, pp. 1367–1403, 2014.

[58] D. Bradley, J. H. Jr., R. Jame, and Z. Xiao, "Place your bets? the market consequences of investment advice on reddit's wallstreetbets," *SSRN Electronic Journal*, 2021.

[59] A. Atkins, M. Niranjan, and E. Gerding, "Financial news predicts stock market volatility better than close price," *The Journal of Finance and Data Science*, vol. 4, no. 2, pp. 120–137, 2018.

[60] P. Kumar and E. Foufoula-Georgiou, "Wavelet analysis for geophysical application," *Reviews of Geophysics*, vol. 35, pp. 385–412, 11 1997.

[61] R. Kronland-Martinet, "The wavelet transform for analysis, synthesis, and processing of speech and music sounds," *Computer Music Journal*, vol. 12, no. 4, pp. 11–20, 1988.

[62] J. B. Ramsey and C. Lampart, "Decomposition of economic relationships by timescale using wavelets," *Macroeconomic Dynamics*, vol. 2, no. 1, pp. 49–71, 1998.

[63] P. M. Crowley, "A guide to wavelets for economists," *Journal of Economic Surveys*, vol. 21, no. 2, pp. 207–267, 2007.

[64] A. Rua and L. C. Nunes, "International comovement of stock market returns: A wavelet analysis," *Journal of Empirical Finance*, vol. 16, no. 4, pp. 632–639, 2009.

[65] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bulletin of the American Meteorological Society*, vol. 79, no. 1, pp. 61 – 78, 1998.

[66] Z. Umar, O. B. Adekoya, J. A. Oliyide, and M. Gubareva, "Media sentiment and short stocks performance during a systemic crisis," *International Review of Financial Analysis*, vol. 78, p. 101896, 2021.

[67] Z. Umar, M. Gubareva, I. Yousaf, and S. Ali, "A tale of company fundamentals vs sentiment driven pricing: The case of gamestop," *Journal of Behavioral and Experimental Finance*, vol. 30, p. 100501, 2021.

[68] Z. Umar, I. Yousaf, and A. Zaremba, "Comovements between heavily shorted stocks during a market squeeze: Lessons from the gamestop trading frenzy," *Research in International Business and Finance*, vol. 58, p. 101453, 2021.

[69] C. Long, B. M. Lucey, and L. Yarovaya, ""i just like the stock" versus "fear and loathing on main street" : The role of reddit sentiment in the GameStop short squeeze," *SSRN Electronic Journal*, 2021.

[70] N. Fusari, R. Jarrow, and S. Lamichhane, "Testing for asset price bubbles using options data," *SSRN Electronic Journal*, p. 51, 11 2020.

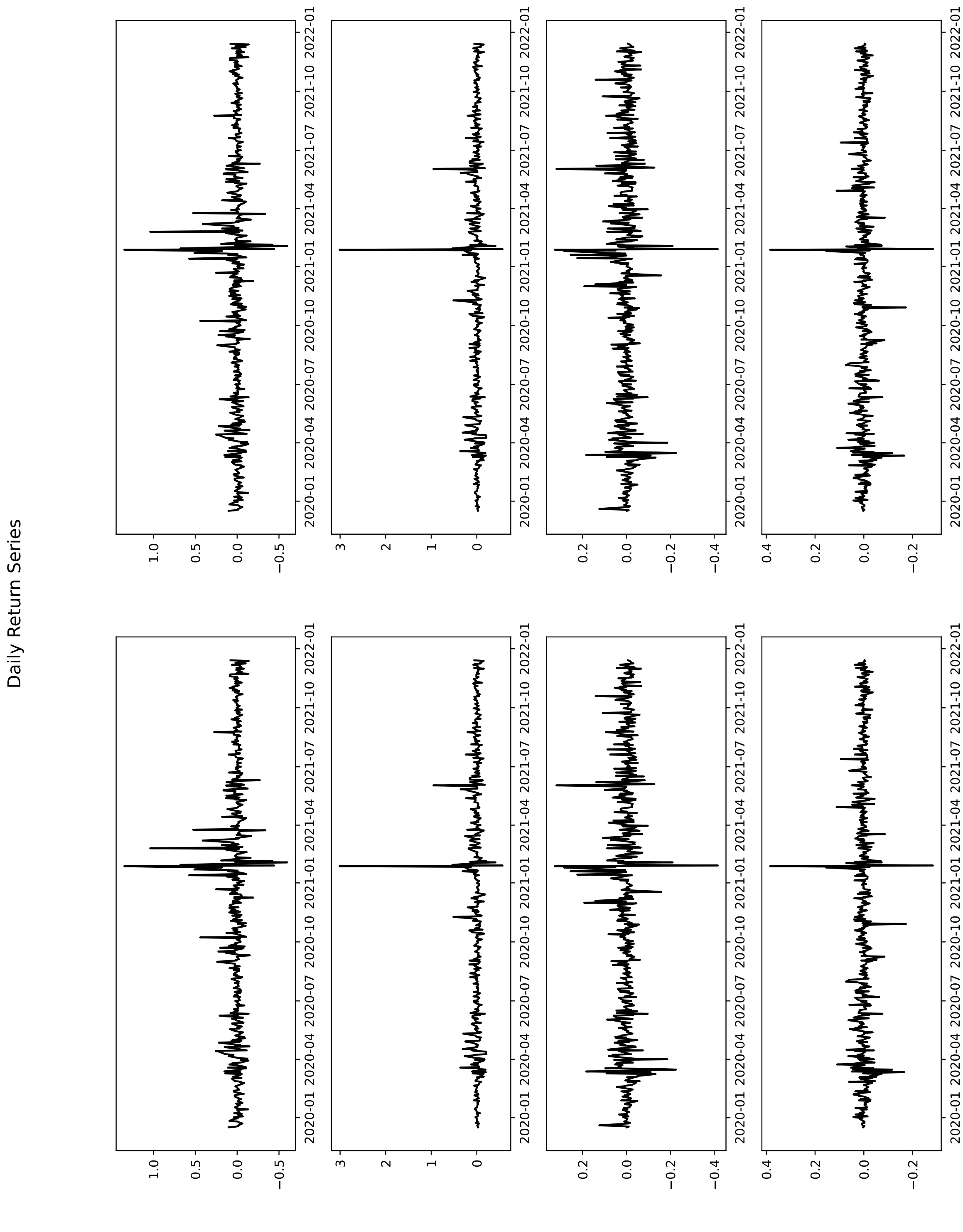
[71] C. M. Jones, A. V. Reed, and W. Waller, "When brokerages restrict retail investors, does the game stop?," *SSRN Electronic Journal*, p. 79, 11 2021.

[72] M. P. Franklin Allen, Eric Nowak and A. Tengulov, "Squeezing shorts through social news platforms," *SSRN Electronic Journal*, p. 73, 8 2021.

[73] G. Baltussen, Z. Da, S. Lammers, and M. Martens, "Hedging demand and market intraday momentum," *Journal of Financial Economics*, vol. 142, no. 1, pp. 377–403, 2021.

[74] R. J. Shiller, "Narrative economics," *The American Economic Review*, vol. 107, no. 4, pp. 967–1004, 2017.

Appendix A



Appendix B

If you need more then one section...

@misctime, author=, title=, year = , month = , howpublished=, note =