Social contagion and asset prices: Reddit's self-organised bull runs*

Valentina Semenova[†1,2,3] and Julian Winkler[‡3,4]

[1]*Department of Mathematics, University of Oxford*
[2]*Oxford Man Institute for Quantitative Finance*
[3]*Institute of New Economic Thinking at the Oxford Martin School*
[4]*Department of Economics, University of Oxford*

Abstract

This paper investigates how social forces can affect asset prices, through an empirical analysis of the beliefs and positions of individuals active on Reddit's WallStreetBets (WSB) forum. The paper proposes a stylized model, incorporating complementarities in asset demand alongside price extrapolation. In doing so, we explain three features of the WSB data: i) users shift their expectations about returns to match those of their peers, ii) returns are predictable and experience reversals, and iii) heterogeneity in investor sentiment can affect markets when popularity is heavy-tailed. We empirically document that sentiments expressed by WSB users about assets' future performances (bullish or bearish) are in part due to sentiments of their peers and past asset returns. The paper also estimates the effect of WSB activity on asset prices, and provides insight into how heterogeneous, idiosyncratic sentiments can impact asset prices. The paper concludes that the unique features of social media, such as its scale and ability to facilitate the spread of information, make it an important driver of market behaviour.

JEL codes: D91, G14, G41.

*Acknowledgements: We are grateful to participants of the Royal Economic Society annual conference, the Alan Turing Institute interest group on economic data science, and the Rebuilding Macroeconomics seminar. We are particularly indebted to Rick Van der Ploeg, J. Doyne Farmer and Xiaowen Dong for their valuable feedback. We also benefitted from insightful comments by Andrei Shleifer, Xavier Gabaix, Steve Bond, Kieran Marray, Paul Tetlock, Laura Veldkamp, David Hirshleifer, Francis DiTraglia, Kevin Sheppard, John Pougué-Biyong, Ilan Strauss, Jangho Yang, François Lafond, Matthias Winkler, José Moran, Farshad Ravasan, Pedro Bordalo, Cars Hommes, Stefan Zohren, Mirta Galesic, and Renaud Lambiotte. We thank Baillie Gifford and the Institute for New Economic Thinking at the Oxford Martin School for funding our work at the University of Oxford.

†valentina.semenova@maths.ox.ac.uk

‡julian.winkler@economics.ox.ac.uk

1 Introduction

In investigating the stock market crash of May 28, 1962, the Securities and Exchange Commission (SEC) found that: 'investor "psychology" being what it is, the increasing decline in one or several issues can easily spread to others. Once the process becomes generally operative, the stage is set for a serious market break' (1963). The SEC concluded that large institutions acted as a balancing force during the collapse. The report pointed at retail traders as the *key players* behind the panic. Over half a century later, we are again confronted with the consequences of investors' social behaviours. As online discussants on Reddit's 'WallStreetBets' (WSB) forum drove up the price of GameStop shares in January, 2021, retail investors regained a spotlight on the (virtual) trading floor. A key difference between 1962 and today is the internet, which offers both a coordination platform on an unprecedented scale, and a new datasource on investor narratives, interactions and psychology.

This paper sets out to reconcile observed behaviours on social media with economic theory by examining the beliefs and positions of individuals active on WSB. We propose a stylized model for how social and psychological forces – information assimilation from peers and extrapolation – can affect asset prices and result in bubbles and volatility. The model's assumptions are validated using the WSB dataset. The data is also used to explore several other aspects of investor behaviour, such as their reaction to market surprises and the passthrough of beliefs to asset demand. In a final empirical exercise, we estimate the effect of WSB activity on asset prices.

Our approach addresses several challenges in the literature. Current research often relies on investor survey data for information on beliefs, and filings (such as 13F filings) to study holdings. These are reported at a fairly low frequency (quarterly or semi-annually), typically cover high net worth individuals or institutional investors (and only their largest asset holdings) and do not explicitly match holdings to beliefs. The WSB dataset allows us to examine small retail investors and observe their reactions to information at a more granular timescale. We are able to track positions and beliefs of the same individuals over time, as well as the information that they are exposed to on the forum. In combination with novel econometric techniques, these rich data allow us to produce estimates for retail investors' reactions to information shared by peers as well as large market moves, and the extent to which WSB discourse has moved asset prices at a daily or weekly frequency.

We begin with an empirical analysis of the data. First, we examine the extent to which individuals trade based on their stated beliefs about an asset's predicted performance. We manually extract positions from screenshots publicised on WSB and sentiments using supervised text analysis. Expressing a positive sentiment about an asset on WSB raises the probability of a long investment in the same asset in the future by over six times. The effect is not symmetric - expressing a negative sentiment raises the probability of a short investment only 2.5 times. Neutral sentiments appear to be highly predictive of long positions. The strong, statistically significant link between the sentiment of a WSB user and their subsequent positions demonstrates the credibility of the discourse on the forum, in terms of compelling users to trade along their stated interests.

We perform a preliminary analysis of how beliefs relate to asset prices. We observe that higher asset returns are associated with high current sentiment, but lower past sentiments. On a day when users on WSB express sentiments that are two times more bullish than bearish about an asset, we would expect to see an excess positive return of approximately 0.2 percentage points (pp) in the asset. However, on the next day, we would expect excess returns to decline by 0.05 pps, constituting a price reversal. We track the link between returns and sentiments in the twenty days surrounding activity on WSB: we find a statistically significant, positive relationship between current sentiments and returns up to seven days in the past, and a negative relationship with returns up to four days in the future. Investors on WSB appear to follow trends in prices, after which returns revert.

Our stylized model sheds light on the relationship between investor beliefs and returns. We consider risk-averse investors who demand assets with uncertain payoffs based on personal valuations, and we assume that there is some exogenous supply of the asset. Knowledge of other investors' demand is useful because it helps investors predict the tendency of the price to go up or down through market clearing. This knowledge complements individual valuations and can impact returns.

In a dynamic setting, we expect our proxy for asset demand by WSB users to drive prices up contemporaneously but depress future returns since they are willing to hold the asset purely due to social pressure. The model also demonstrates how heterogeneous, idiosyncratic sentiments can survive aggregation and impact asset prices if the mechanism by which investors are exposed to each others' choices exhibits a heavy-tailed pattern. For example, some submissions on WSB garner tens of thousands of comments, even though most receive fewer than ten. This is a promising avenue for further research that emphasizes the role of 'virality' in social media content, like the one which underpinned the infamous GameStop short-squeeze.

The second part of our paper presents an empirical estimation of our model's parameters, focusing on social contagion and price extrapolation among users of WSB. Our main goal is to quantify the extent to which expressed sentiments are influenced by so-called 'peer effects'. To accomplish this, we test how peer sentiment impacts investor decision-making in two ways. First, we select individuals who express sentiments about an asset multiple times and observe how peers discuss the same asset in between. We use historic peer sentiment as an Instrumental Variable (IV), which mitigates the common shock problem. This approach is inspired by the peer effects in the classroom literature, which gauges future student performance based on entry exams (Duflo et al. 2011). Second, we leverage the WSB network of interactions to identify the information to which an individual investor has been exposed. The network links an older submission about an asset to a new submission if the author of the new submission comments on the older submission. To estimate a parameter for social contagion, we regress the sentiment expressed by the new submission on the average sentiment of older, linked submissions. We use the timing of our IVs to control for common shocks and instrument the sentiment of linked submissions to control for an author's endogenous choice to comment.

In both approaches, exogenous variation in average peer sentiment is a statistically significant predictor for the change in author sentiment. This finding suggests that retail investors experience investment complementarities and adapt their strategies based on those of their peers. The instrumental variable (IV) results reveal that when the odds of peers expressing bullish over bearish sentiments double, the odds of a given user expressing bullish over bearish sentiment increase by an average of 14%. Although the role of narratives in investor decision-making has been extensively discussed in the literature, this work is, to the best of our knowledge, the first to document a relationship between an investor's sentiment and that of their peers outside a controlled experimental setup (Bursztyn et al. 2014).

This empirical exercise sheds light on several phenomena regarding how investors respond to various signals. First, we confirm that retail investors on WSB tend to be trend-followers. Specifically, a log-return of 0.1 on a given day increases the probability of a user posting a bullish over bearish post by twenty percent. Second, we explore how retail investor sentiments respond to market surprises, defined as log-returns that are two standard deviations above or below its monthly average. Interestingly, the effect is negative for both positive and negative market surprises, but the positive surprise has a low level of statistical significance. This finding suggests that downside panic can spread quickly among investors. Finally, our experiment highlights the role that reinforcement between peer sentiments and market performance plays.

In the final section, we synthesize our model and empirical observations of WSB users to measure their impact on returns. Our main strategy centres on predicting variation in sentiments among WSB users unrelated to current price changes, leveraging the strong temporal persistence of sentiments to specific assets due to the peer effect channel. Our estimates are both statistically and economically significant in predicting changes in weekly average log returns, providing evidence for a relationship between social dynamics, as proxied by WSB conversations, and financial markets. Overall, our analysis underscores the importance of considering the social aspect of retail trading when attempting to understand the dynamics of financial markets.

An important question that remains unanswered is how to explain socially-driven price run-ups, such as the GameStop short squeeze. To tackle this issue, we extend our model to account for idiosyncratic social shocks in sentiments that cannot be explained by recent news and stock performance, and are distinct from our overall sentiment metric. Specifically, we leverage the framework of Gabaix & Koijen (2020, 2021) for granular instrumental variables, and we propose that 'granular' social shocks, i.e., those that receive a large social following, will not be averaged out across all investor sentiments and can have an outsized impact on returns. Importantly, as with other online content, our data is heavy-tailed, with only a few submissions receiving a large following while most go unnoticed.

Our study introduces a granular instrumental variable that captures the impact of idiosyncratic social shocks on market returns. To construct this variable, we compare the average sentiment of all submissions to the sentiment of only the most popular submissions. Our analysis reveals a statistically significant relationship between our predicted social shock and future returns. Specifically, a doubling in the odds of a very popular submission expressing bullish over bearish sentiments results in an average increase in returns of 0.007 the following week. However, the impact of negative idiosyncratic social shocks is much greater than positive ones. Our findings help shed light on the dynamics behind the slow build-up and rapid decline of popular stocks on WSB, as well as other market panics.

Related literature The economic interest in asset mispricing has a long history, with examples dating back to Tulipmania in the Netherlands in the 17$^{\text{th}}$ century (Garber 1989). Since then, numerous frameworks have been proposed to explain the gradual increase and sudden drop in financial assets. These frameworks include psychological models, such as diagnostic expectations (Bordalo et al. 2021), the spread of information (Veldkamp 2006), strategic complementarities (Hellwig & Veldkamp 2009, Zenou 2016), and extrapolation (Glaeser & Nathanson 2017). In this paper, we draw upon several approaches in the literature to frame our model, which emphasizes extrapolation and complementarity in asset demand among investors.

A parallel strand of the literature highlights the importance of peers and narratives in shaping investor perspectives. In his seminal work, Shiller (1984) discusses the excess volatility in stock prices relative to dividends. Since then, 'narrative economics' has played an increasingly important role in our understanding of investor decision-making and market moves (Shiller 2005, 2014, 2017, Banerjee et al. 2013, Hirshleifer 2020). Our work leverages new data to provide fresh empirical evidence of how heuristics and information gained from peers can affect investor decision-making and lead to asset mispricing.

Several studies in the peer effects literature leverage naturally occurring variation in peers for their identification strategy. An area which pioneered many of these techniques investigates peer effects in the classroom (see Epple & Romano (2011), Sacerdote (2011) for a general overview, and Duflo et al. (2011) for a prominent example). Social networks are also an active area of study (see Bramoullé et al. (2020) for a recent review). Bursztyn et al. (2014) perform a field experiment with a financial brokerage in Brazil, where they study investment decisions made by peer pairs: the peers are offered a 'high stakes' investment opportunity (minimum investments were R$2,000 – around 50% of the median investor's monthly income) in a certain order to identify the effects of 'social learning' and 'social utility' in financial decision-making.

Other related work investigates the diffusion of micro-finance decisions in a social network (Banerjee et al. 2013), the effect of peers on risk taking (Lahno & Serra-Garcia 2015), and the effect of social networks on saving (Breza & Chandrasekhar 2019). By studying a broader set of investors in a natural experiment, our research question is similar to Pool et al. (2015), who demonstrate that socially connected fund managers appear to hold similar stocks. The present paper highlights how to transfer well-established techniques from this literature to social media data. We also leverage novel empirical techniques, such as granular instrumental variables, designed specifically to tackle confounding issues in financial markets (Gabaix & Koijen 2020, 2021).

There exist studies on the interplay between online forums and financial markets, as well as the spread of information in social networks (Sabherwal et al. 2011, Tetlock 2007). Our work differs from studies focusing on the spread of information through friend networks, such as Aral et al. (2009), Aral & Nicolaides (2017), as Reddit users are anonymous without explicitly defined friendship links. The anonymity within Reddit is crucial to the prominence of WSB: in contrast to the exercise in Banerjee et al. (2013), where information is transmitted via friendship networks, the mechanism by which information dissipates on WSB is much closer to the homogenous mixing conditions popular in traditional epidemiological models, and therefore closer in spirit to Banerjee (1993). Our work complements studies that focus on identifying one direct relationship between social activity and assets, for example Antweiler & Frank (2004), Kumar & Lee (2006), Chen et al. (2014). Cookson & Niessner (2020) also use social media data to understand investor social dynamics and demonstrate an impact to the financial markets, but focus on disagreement among investors.

Road map We present our results in five sections. The following section comprehensively describes the data source and relevant variables. Section 3 presents a model for price dynamics in the presence of information sharing among investors. Section 4 presents empirical evidence for our proposed investor dynamics. Section 5 empirically evaluates the effect of retail investors on financial markets. Section 6 concludes.

2 What is *WallStreetBets*?

Reddit, launched in 2005, is a social news aggregation, web content rating, and discussion website. It was ranked as the 8th most visited site globally in November 2022,[1] with over 430 million anonymous users by the end of 2019.[2] The website's contents are self-organised by subject into smaller sub-forums, 'subreddits', which discuss a unique, central topic.

Structure of WSB Within subreddits, users publish titled posts (called 'submissions'), typically accompanied with a body of text or a link to an external website. These submissions can be commented and 'upvoted' or 'downvoted' by other users. A ranking algorithm raises the visibility of a submission with the amount of upvotes it receives, but lowers it with age. Therefore, the first submissions that visitors see are i) highly upvoted, and ii) recent, with the precise algorithm considered private intellectual property and discussed further in Appendix A.1.[3] Comments on a submission, visible to anyone, are subject to a similar scoring system, and can, themselves, be commented on.

Features The WSB subreddit was created on January 31, 2012, and reached one million followers in March 2020.[4] As per a Google survey from 2016, the majority of WSB users are 'young, male, students that are inexperienced investors utilizing real money (not paper trading); most users have four figures in their trading account'.[5] Individuals on the forum

[1] https://www.statista.com/statistics/1201880/most-visited-websites-worldwide/
[2] https://redditblog.com/2019/12/04/reddits-2019-year-in-review/
[3] https://www.reddit.com/r/help/comments/717686/order_of_posts/
[4] https://subredditstats.com/r/wallstreetbets
[5] https://andriymulyar.com/blog/how-a-subreddit-made-millions-from-covid19

discuss and express their sentiments about stock-related news. In addition to market discussions, there is ample evidence of users pursuing the investment strategies encouraged in WSB conversations. Users post screenshots of their investment gains and losses, which subreddit moderators are encouraged to verify – a dynamic reminiscent of Shiller's (2005) description of an asset bubble. The discussions are whimsical, but mostly investment-focused.

Available data We downloaded WSB data using the PushShift API.[6] PushShift records all comment and submission data at the time of creation. The full dataset consists of two parts. The first is a total of 1.4 million submissions, with their authors, titles, text and timestamps. The second is comprised of 16.5 million comments, with their authors, text, timestamp, and the identifier of the parent comment or submission. Submission and comment numbers have grown exponentially since 2015 – Figure 7 in Appendix A.1 displays the forum's exponential growth.

Our dataset spans January, 2012 to July, 2020. Importantly, it does not include the events of the 2021 GameStop (GME) short squeeze. The decision to focus on this timeframe is intentional: before the GME short squeeze, WSB received less attention from institutional investors, as well as less bot-activity. As such, our sample tracks retail investor discussions more precisely, without systematic external influence. Furthermore, ample research has emerged focusing exclusively on the GameStop short squeeze, whereas our goal is to characterise investor behaviour, rather than examine a single event.

Identifying assets The following sections predominantly rely on submissions for text data, since they are substantially richer than individual comments. Comments are used to trace interactions between discussants. In order to understand how users discuss specific assets, we extract mentions of *tickers* from the WSB submissions' text data. A ticker is a short combination of letters, used to identify an asset on trading platforms. For example, 'AAPL' refers to shares in Apple, Inc. Appendix A.2 documents how tickers are extracted from submissions. Table 8 in Appendix A.2 displays the twenty tickers that feature most prominently in WSB conversations up to July, 2020. These are typically stocks in technology firms, such as AMD or FB. A handful of Exchange Traded Funds (ETFs) are also present, notably the S&P 500 (SPY) and a leveraged gold ETF (JNUG).

A small fraction of the 4,650 tickers we extract dominate the discourse on WSB: 90% of tickers are mentioned fewer than 31 times, and more than 60% are mentioned fewer than five times. Appendix A.2 documents the heavy-tailed nature of ticker discussions. In total, we are left with 111,765 submissions that mention one, unique ticker and were posted before July 1st, 2020. These submissions have 1.9 million comments in total.

Sentiment model In addition to extracting tickers, we gauge whether submissions express an expectation for an asset's future price to rise, the *bullish* case, to fall, the *bearish* case, or to remain unpredictable/stable, the *neutral* case. We identify sentiment using a supervised-learning approach, with a hand-labeled dataset of almost five thousand submissions for

[6] https://pushshift.io/

training, validation and testing (Araci 2019). The sentiment model outputs a probability for each sentiment category, achieving 70% accuracy in categorising the manually labeled test set. Appendix A.3 discusses details of this Natural Language Processing (NLP) model and the distribution of labels.

Key sentiment variable The sentiment classifier assigns three probability scores to each submission about a ticker: the probability of a submission being bullish, $P(\phi = +1)$, bearish, $P(\phi = -1)$, neutral, $P(\phi = 0)$. The probabilities sum to one. At the time t when an author i posts about asset j, we use the probability scores above to calculate a continuous sentiment score between $(-\infty, \infty)$:

$$\Phi_{i,j,t} = \frac{1}{2}\log\left(\frac{\mathrm{P}(\phi_{i,j,t} = +1)}{\mathrm{P}(\phi_{i,j,t} = -1)}\right). \tag{1}$$

Submissions labeled as bullish ($\mathrm{P}(\phi = +1) = 1$), or bearish ($\mathrm{P}(\phi = +1) = 1$), are set to $\mathrm{P}(\phi = +1) = 0.98$, or $\mathrm{P}(\phi = -1) = 0.98$, to retrieve a finite value for the log-odds. We also extract three categorical variables (bullish, bearish, neutral) which are encoded with a one if the label received the highest probability from our classifier: the categorical variable $\phi^{+1}_{i,j,t}$ will be equal to one is author i's post about asset j at time t is categorised as bullish; $\phi^{0}_{i,j,t}$ and $\phi^{-1}_{i,j,t}$ will be zero. We leverage these variables to investigate investor sentiment throughout the paper. Appendix A.3 shows the distribution of our key sentiment variable $\Phi_{i,j,t}$.

2.1 Isn't all of this just talk?

Table 1: Follow-through on WSB Advice

	Dependent variable: Position in Asset j of Author i	
	$B_{i,j}$ - categorical	
	(1)	(2)
$\Phi_{i,j}$	1.50 (0.20) ***	
$\phi^{-1}_{i,j}$		-0.97 (0.29) ***
$\phi^{0}_{i,j}$		0.66 (0.21) ***
$\phi^{+1}_{i,j}$		1.84 (0.27) ***
Observations	278	278
Pseudo-R^2	0.13	0.17

Notes: This table presents estimated log-odds coefficients for two logit models for the relationship between the sentiment expressed by user i about asset j and the subsequent long/short position the user reports (see Eq. 2 for a further description). Sentiment estimates are presented in two ways: (1) the continuous log-odds of the author expressing positive over negative sentiment $\Phi_{i,j}$, and (2) as a categorical variable where $\phi^{-1}_{i,j}$ corresponds to the expression of negative sentiment, $\phi^{0}_{i,j}$ - neutral, $\phi^{+1}_{i,j}$ - positive.
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Why should we care about the sentiments people express about assets online? Anecdotally, the GameStop short squeeze demonstrated that the online discussions on the WSB forum have impact on assets. However, this does not constitute evidence of the fact that people follow through on the investment strategies they discuss online systematically.

To address the concern that WSB sentiment data has limited impact on investment decisions, we utilize screenshots that users post of their investment positions to test whether they follow through on their expressed sentiments. We extract approximately 9,000 images from WSB – we focus only on image-related URLs (such as ones with the domain name 'imgur', an image-hosting site) mentioned in posts of authors who had previously posted about a ticker. We hand-annotate a third of the images. Specifically, we manually annotate the image, if it is a position screenshot, with i) the tickers in the screenshot and ii) the positions (long or short) the author displays. The position taken by author i in asset j, $B_{i,j}$, are annotated as +1 if the author is long in the asset, and -1 if the author is short.

We note that the sample of screenshots is biased. Authors on WSB are socially incentivized to share extreme losses or gains. We, therefore, observe relatively few positions, as compared to sentiments. The positions data is also skewed towards long positions, which is consistent with the skew towards bullish sentiments on the forum. However, despite these shortcomings, the positions provide sufficient variety in investment strategies to test whether people trade on their expressed sentiments on WSB.

We match the ticker screenshot to a submission posted before/simultaneously with that screenshot by the same author and about the same ticker. We regress the most recently expressed sentiment by author i about asset j (our key sentiment variable $\Phi_{i,j}$) on the log-odds of the position $B_{i,j}$ extracted from their screenshot being long versus short:

$$\log\left(\frac{\mathrm{P}(B_{i,j}=+1)}{\mathrm{P}(B_{i,j}=-1)}\right)=\lambda^{s}\Phi_{i,j}+u^{p}_{i,j,t}, \tag{2}$$

where λ^s measures the pass-through rate of sentiment into eventual investment positions. In an alternative formulation, we represent the past sentiment as three categorical variables: $\phi^{-1}_{i,j}$, $\phi^{0}_{i,j}$, $\phi^{+1}_{i,j}$, which take on a value of one if the author's sentiment is labeled short, neutral or long, respectively (and a value of zero otherwise).

Results Table 1 presents the coefficients of past sentiments regressed on future positions, estimated using a Logistic regression. We observe that an author's sentiment is highly correlated to their subsequent holdings of the stock. Let us consider the results in column (2) – an author creating a bullish post about an asset raises the probability of a long versus short investment by over six times. The positions data gives us confidence that investors do trade based on their discussions and expressed sentiments.

2.2 Predicting stock returns with WSB sentiments

Does activity on WSB influence stock market returns? We run a set of simple exercises to motivate our future analysis. We average the sentiment characteristics in Eq. 1 by stock j and trading day t, denoting these mean sentiments by $\bar{\Phi}_{j,t}$. We merge these daily sentiment observations with US common stock returns reported by CRSP (detailed in Appendix A.4), and transform the reported returns into log returns.

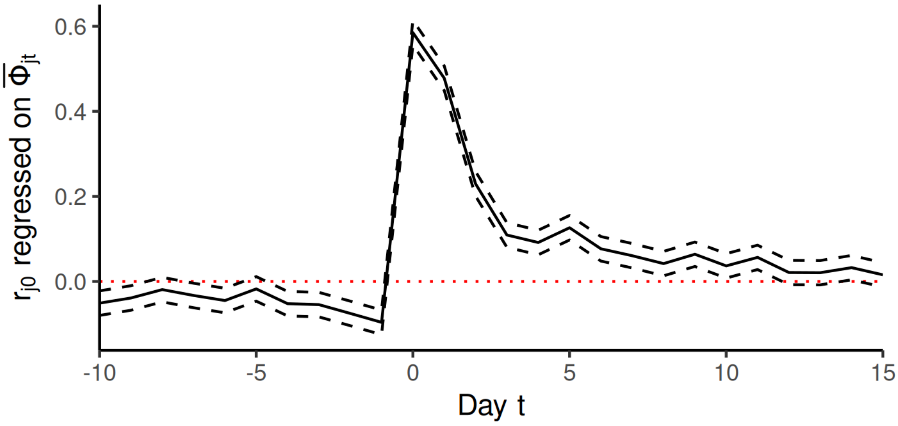


Figure 1: **Return correlation with WSB sentiments;** Daily stock log returns are regressed on daily average sentiments expressed in submissions on that stock on WSB, where sentiments are lagged by -10 to 15 days – the estimated relationship is described in Eq. 3. A lag of -5 implies that sentiments precede the returns observation by five days. The point estimate for the correlation is plotted by the solid black line, with the 99% confidence interval in dashed black lines. A correlation of zero is highlighted by the dotted red line. All variables are demeaned by their daily average. Return correlation with past WSB sentiments is negative, but strongly positive with current and future sentiments.

Table 2: Stock returns versus WSB characteristics

	Dependent variable:	
	$r_{j,t}$	
	(1)	(2)
$\bar{\Phi}_{j,t}$	0.60*** (0.04)	
$\bar{\Phi}_{j,t-1}$	−0.16*** (0.02)	−0.07*** (0.02)
$r_{j,t-1}$		−0.06*** (0.004)
$\bar{\Phi}_{j,t-1} \times r_{j,t-1}$		0.01 (0.01)
Day FE	Yes	Yes
Observations	8,287,639	8,287,639
R^2	0.0004	0.003

Notes: This table presents the OLS estimates for the relationship between stock log returns, $r_{j,t}$ and average expressed sentiment on WSB, $\bar{\Phi}_{j,t}$ and $\bar{\Phi}_{j,t-1}$. t represents time in days. The regression highlights the existence of a positive relationship between current sentiments and current returns, and a negative one between current returns and past sentiments. The negative relationship persists when controlling for previous day returns $r_{j,t-1}$ in Column 2. Accompanying standard errors, displayed in brackets, are clustered at the stock level, and calculated in the manner of MacKinnon & White (1985).
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

We first consider a regression of log returns on mean daily sentiment,

$$r_{j,t} = \lambda_{1,T}\bar{\Phi}_{j,t+T} + \eta_t^r + v_{j,t}^r, \tag{3}$$

where $v_{j,t}^r$ is a residual, η_t^r is a daily fixed effect, and T denotes a lag varying from -10 to 15 days. The OLS estimates for coefficients $\lambda_{1,T}$ describe how WSB sentiments are temporally related to stock returns.

Subsequently, we regress daily log returns on current mean sentiments, as well as previous day sentiments:

$$r_{j,t} = \lambda_1 \bar{\Phi}_{j,t} + \lambda_2 \bar{\Phi}_{j,t-1} + \eta_t^r + v_{j,t}^r, \tag{4}$$

where η_t^r is a daily fixed effect, $v_{j,t}^r$ an error term, and λ_1, λ_2 our coefficients of interest. This specification gives a sense for the dynamic properties of WSB sentiments – leading up to a trade day, plus their response on that day.

Results Figure 1 plots the OLS estimates for $\lambda_{i,T}$ in Eq. 3 as a function of lag T. Generally, past sentiments appear negatively related with current returns, although the effect is small, and not highly significant beyond two lags. Current sentiments are strongly correlated with current returns, and this effect persists for about five days, before dissipating. This implies that a large return in an asset today will have a persistent impact on investor sentiment for five days into the future. Investor sentiments do not anticipate future returns, but rather follow the trend.

Table 2 reports OLS estimates for coefficients from Eq. 4 in Column 1. Returns again relate positively to contemporaneous sentiment, and negatively with previous day sentiments. Both of these are statistically significant at the 1% level. However, the implied effects are relatively small; on an average day, returns are 0.1 log points lower if sentiments expressed on the previous day are twice more likely to be bullish than bearish.

In Column 2, we estimate Eq. 4 with the interaction between lagged, average sentiment and lagged returns, to capture a non-linearity for sentiments in stocks that garner exceedingly high amounts of attention. The slight negative relationship between current and past returns could potentially confound the effect of past sentiment, as seen in the smaller coefficient for lagged sentiments. However, there is no clear evidence that the interaction between sentiments and outsized returns produce a significant effect on subsequent returns.

2.3 Motivation

If Table 1 is to be believed, the sentiments expressed on WSB induce trading activity. However, the negative relationship between current returns and past sentiment reported in Table 2 is puzzling in that regard, and would suggest that the authors of those submissions erred in their assessments.

A central argument of our paper is that the correlations in Table 2 are not a manifestation of (erroneous) information spreading on WSB. Rather, the trading patterns of these retail investors are responsible for reversals in prices as they seeks to find and follow highly risky strategies. The following section builds a hypothesis on the emergence of bubble-like dynamics as a function of social contagion in investor strategies, whereby return expectations are based, in part, on experiences of peers.

3 Social dynamics and asset prices

What motivates investors to share trading advice online? Some seasoned traders might argue that one can only have an investment edge while others are unaware of your strategy. We rationalize our observed online interactions through an asset demand model by incorporating information complementarities in investment decisions. This gives rise to *social contagion* in asset demand; investors buy the asset because others do as well, irrespective of their personal beliefs. The model explores the patterns of returns when this behaviour is present.

Our proposed mechanism is motivated in part by a recent literature that studies diagnostic expectations (Bordalo et al. 2021). Investors trade on the momentum of the stock price, against a supply of shares provided by noise traders. Our inclusion of a social component subsequently induces persistence in asset demand over time, which leads to reversal in future returns.

3.1 General setup

We analyse price for one asset traded by N investors, indexed by i. Each investor derives CARA utility from consuming c, $U_i(c_i) = \exp(-\gamma c_i)$, where γ is the constant absolute rate of risk aversion. We do not include any discounting in their decision-making, but assume they evaluate an asset according to a log-normal distributed value $v = \log(V)$ with expectation $\mathbb{E}_i(v)$ and variance $\sigma_i^2(v)$.

To build intuition, we drop the time subscript and initially present a static version of the model. Investor i purchases ϕ_i shares at the current market log-price p to optimise the mean-variance objective function

$$\mathcal{L}(\phi_i) = [\mathbb{E}_i(v) - p]\phi_i - \frac{\gamma}{2}\sigma_i^2(v)\phi_i^2, \tag{5}$$

$$\Rightarrow \phi_i^* = \frac{\mathbb{E}_i(v) - p}{\gamma\sigma_i^2(v)}, \tag{6}$$

where an asterisk denotes the value that maximises objective $\mathcal{L}$. In this way, we distinguish between beliefs about value $\mathbb{E}_i(v)$ from investor i's decision to buy amount ϕ_i. Eq. 6 yields a familiar expression for asset demand in equilibrium: namely a ratio of expected net returns over their variance.

Simple price equilibrium We assume that asset supply originates from noise traders, as in Bordalo et al. (2021), to find a price equilibrium. In effect, equilibrium demands $\sum \phi_i^*$ sum to exogenous supply S. Assuming fixed uncertainty $\sigma^2 = \sigma_i^2(v)$ and averaging expected values $\mathbb{E}(v) = 1/N \sum \mathbb{E}_i(v)$, we can re-arrange Eq. 6 to yield the following expression for the market-clearing price:

$$p = \mathbb{E}(v) - \frac{S}{N}\gamma\sigma^2. \tag{7}$$

Eq. 7 accounts for the price level by investor's average expected value of the asset, in addition to their ability to absorb the exogenous level of assets supplied. This ability depends on the depth of the investor pool – reflected by the number of investors N – as well as their risk appetite $\gamma\sigma^2$. In this simple market, the price increases with expected value, and decreases with supply.

Goals The first goal of this section is to demonstrate the extent to which asset demand between investors is a strategic complement. The model we consider imposes no cost on investors' information acquisition, which means they anticipate the price of the asset to be high when asset demand by other investors is high. A long literature investigates the link between subjective asset price beliefs and prices, and our contribution is to include a term for social learning.

The current expression for price ignores heterogeneity and aggregation. Eq. 7 mimics a single representative investor of some size N. However, strategic interactions may give rise to feedback loops in investor demands that do not average out. We extend the model to consider investor heterogeneity. Investors central in a social network are 'indirectly' large by their ability to sway others. We make use of this in our empirical study in Section 5.

The final goal is to investigate the equilibrium price in a dynamic setting where investors are influenced not only by information from peers, but also by other psychological factors - namely extrapolation (Barberis et al. 2018, Glaeser & Nathanson 2017). The model produces several predictions about the behaviour of asset prices in the presence of such investors, which we empirically validate later in the paper. We highlight the empirical challenge in a scenario where investor sentiments respond to real-time market events, but point out that the persistence of a social signal may produce variation in asset demand orthogonal to current returns. We use this fact in Section 5 to gauge the extent WSB users exert market impact.

3.2 Static model with investment complementarities

We study the role of complementary investment decisions. To that end, the model operates in two stages. In the first stage, investors build their expectation for the asset's value, using observed signals from their peers and their expectation of the market-clearing price as a function of the expected, and as of yet hidden, supply shock. In the second stage, the asset supply shock is revealed, and investors execute their trades according to their demand curve.

Let's consider an investor i who has some individual value for the asset v and some expectation for others' demand. Investor i's maximised payoff from Eq. 5 is

$$\mathcal{L}(\phi_i^*) = \frac{\mathbb{E}_i^2(v-p)}{2\gamma \mathbb{E}_i(v-p)^2} \tag{8}$$

$$= \frac{1}{2}\mathbb{E}_i(v-p)\phi_i^* \tag{9}$$

$$= \frac{1}{2}[\mathbb{E}_i(v) - \mathbb{E}(v)]\phi_i^* + \frac{\gamma\sigma^2}{2}\left(\frac{1}{N}\sum_j^N \phi_j^*\right)\phi_i^*, \tag{10}$$

where $\mathbb{E}(v) = 1/N \sum \mathbb{E}_i(v)$ as before. Here, investors base their price expectations on the simple equilibrium in Eq. 7, but use their personal expectations and constant uncertainty σ^2 to forecast price in the second stage.

Eq. 10 demonstrates that the investor's payoff depends on their peers in two regards. First, payoffs increase to the degree that the investor in question expects to outperform others, in terms of the value they realise in the asset. This is seen in the first component, by which buying(selling) the asset increases the payoff to the extent that i's expected value $\mathbb{E}_i(v)$ is higher(lower) than that of their peers. Second, the payoff increase by the average optimal asset demand of all investors in the economy.

Complementarity in asset demand This asset demand model predicts that social interactions – knowledge of other's asset purchases – is a significant component of investors' welfare *in expectation*. Eq. 10 is a well-known formulation for strategic interactions between agents acting under quadratic loss (Zenou 2016). Deriving Eq. 10 with respect to two investors' demands reveals their strategic complementarity:

$$\frac{d^2\mathcal{L}(\phi_i^*)}{d\phi_j^* d\phi_i^*} = \frac{\gamma\sigma^2}{2N} > 0. \tag{11}$$

The emergence of strategic complementarities is due to a crowding effect that investors have on price. The higher asset demand by other investors, the higher the realised price will turn out to be. Before the value of the asset is revealed, investors are motivated to gauge demand by others to better estimate what the price will be, in excess of their personal valuation.

The unweighted average of peer sentiment in Eq. 10 emerges because we did not provide a specific mechanism by which information about asset demand is transmitted. The acquisition of information under some cost to the investor is an interesting extension, although already studied by Hellwig & Veldkamp (2009). Their study offers more rigorous insight into the manifestation of strategic complementarities, as well as the emergence of multiple equilibria, when investors seek to learn about the underlying price from a set of possible signals.

Aggregating heterogeneous demand An investor's demand may have some idiosyncratic preference which is not captured by common factors. For example, an investor may place particular confidence in products he enjoys using, or admire the corporate strategy of certain company leaders. Such sentiments would manifest as idiosyncratic, heterogeneous investor demand.

We re-write Eq. 6 to incorporate some unexplained variation in the beliefs of investor i:

$$\phi_i^* = \frac{\mathbb{E}_i(v) - p + e_i}{\gamma\sigma^2}, \tag{12}$$

$$\Rightarrow p = \mathbb{E}(v) - \frac{S}{N}\gamma\sigma^2 + \frac{1}{N}\sum_{i=1}^{N} e_i, \tag{13}$$

where e_i is the idiosyncratic component of i' asset demand. Consistently with our earlier exercise, investors use personal expectations and constant uncertainty σ^2 to forecast price. This generalises the price equilibrium in Eq. 7 to include the average idiosyncratic sentiment shocks of all investors. Under the assumption that these shocks have a finite variance and mean zero, they should average out to zero by the Central Limit Theorem.

Granular shocks The choice to include idiosyncratic shocks appears to be a trivial extension of the model when the valuations of all investors are weighted equally. However, consider a scenario where some investors have different levels of importance s_i for aggregate demand:

$$S = \sum_{i=1}^{N} s_i \phi_i^*, \tag{14}$$

$$\Rightarrow p = \sum_{i=1}^{N} s_i \mathbb{E}_i(v) - \gamma\sigma^2 S + \sum_{i=1}^{N} s_i e_i, \tag{15}$$

where weights $\sum_i s_i = 1$ for demand to equal supply. This could be the case for several reasons: investors could have different amount of capital or, of greater interest to this paper, some investors may have *more sway* in forming public opinion than others. The importance of key players in a social context has been explored in several economic settings – see Zenou (2016) for a thorough overview. Typically, the most central nodes in a social network have the ability to quickly diffuse information and, therefore, have a high influence on others.

We justify this weighting scheme by the fact that certain users on WSB have a disproportionate effect in shaping the broader discourse. Indeed, WSB is structured to promote viral content, and we would expected a consensus to be formed by key players – or, rather, around key submissions. If the distribution of importance does not have a finite variance – i.e. it is 'heavy-tailed' – then the idiosyncratic shocks would not average out to zero.

3.3 Price dynamics with peer effects

To study the joint dynamics of an asset's price and demand by social investors, we treat aggregate asset demand $\phi = \sum_i \phi_i$ and log-price p as state variables for a dynamic system, indexed by time t. In doing so, we assume that cumulative demand ϕ_t reflects a difference between individual valuations of the asset and the price. We distinguish between two independent components of individual valuations: the private signal of individuals $g(b_{i,t})$ and the signal individuals draw from observations of peers $f(\phi_{i,t})$. Aggregate asset demand and price are

$$\phi_t = \frac{\mathbb{E}_t[g(b_{i,t})] + \mathbb{E}_t[f(\phi_{i,t})] - p_t}{\gamma\sigma^2}, \tag{16}$$

$$p_t = \mathbb{E}_t[g(b_{i,t})] + \mathbb{E}_t[f(\phi_{i,t})] - \frac{S_t}{N}\gamma\sigma^2. \tag{17}$$

Cumulative demand is, therefore, the difference between individual valuations and the market-clearing price, normalized by risk-aversion. The market-clearing price, on the other hand, is the difference between individual valuations and the rate of asset supply, normalized by the number of investors N and their risk appetite $\gamma\sigma^2$, similar to Eq. 7.

A focus of this paper is the relationship between valuations $g(b_{i,t})$ and the social component $f(\phi_{i,t})$. Studies in behavioural finance suggest different expectation formation mechanisms that ultimately deviate from rational expectations (Barberis et al. 2018, Bordalo et al. 2020, 2021). We combine two such features in Assumptions 1 and 2 to propose a testable structure for $\mathbb{E}_t[f(\phi_{i,t})]$ and $\mathbb{E}_t[g(b_{i,t})]$.

Persistent demand The mechanism by which past demand enters current asset demand is by the complementarity in investor payoffs. Investor i's payoff to holding the asset is assumed to increase linearly in average asset demand by others.

Assumption 1. *Social investor i's expectation of future returns is linearly increasing in average asset demand by others: $f(\phi_{i,t}) = \alpha\phi_{t-1}$, where $\phi_{t-1} = 1/N\sum_i \phi_{i,t-1}$ is average asset demand.*

Assumption 1 is in line with the finding that our asset demand model produces strategic complementarities in investor asset demands in Eq. 11. Several extensions of the simple formulation are possible to account for greater complexities in social interactions. It can, for example, be extended to take into account *key players* (Zenou 2016) through changing the way that people weight the demand of others $\phi_{i,t-1}$ in the sum $1/N\sum_i \phi_{i,t-1}$ to $\sum_i s_i\phi_{i,t-1}$, where s_i captures the influence of player i and $\sum_i s_i = 1$.

In more complex settings, we can consider the unique complementarities between connected individuals as described in Zenou (2016), clusters of investors (Bouchaud & Potters 2003), or alternative information spreading / individual targeting models, which have received attention in the recent literature (Galeotti et al. 2020). The added complexity would affect aggregate demand through the expectations of other's demand ϕ_{t-1}.

Mechanical Extrapolation We assume that investors partially trade on the momentum of the asset's price, which Bordalo et al. (2021) term 'mechanical extrapolation'. The functional form of $\mathbb{E}_t[g(b_i)]$ is specified in Assumption 2.

Assumption 2 (Mechanical Extrapolation)**.** *The average investor projects past price increases* *into the future using the updating rule:*

$$\mathbb{E}_t[g(b_{i,t})] = p_t + \beta(p_t - p_{t-1}), \tag{18}$$

where β captures a fixed degree of price extrapolation.

Mechanical extrapolation is our preferred way to introduce a relationship between prices and demand (Barberis et al. 2018). A model with mechanical extrapolation has several shortcomings, one of which is the inability to relate expectation updates to psychological underpinnings. However, our assumption is justified by our empirical work in Section 4.2 which demonstrates that individuals update their outlook based on recent asset returns.

System for price and demand Combining Assumptions 1 and 2 into Eqs. 16-17 yields demands and returns $r_t = p_t - p_{t-1}$:

$$\phi_t = \frac{\alpha\phi_{t-1} + \beta r_t}{\gamma\sigma^2}, \tag{19}$$

$$r_t = -\frac{\alpha}{\beta}\phi_{t-1} + \frac{S_t\gamma\sigma^2}{\beta N}. \tag{20}$$

In this scenario, asset demand and returns are determined simultaneously. The first mechanism is through market clearing, where demand has to adjust to supply. The second is the adjustment of the expected value for the asset to the realised return through β and the social signal through α.

As a result, returns are accounted for by current and past asset demand:

$$r_t = \frac{\gamma\sigma^2}{\beta}\phi_t - \frac{\alpha}{\beta}\phi_{t-1}, \tag{21}$$

since supply must equal demand at time t. This equation uncovers several important mechanisms at play. Returns are related positively to current demand ϕ_t through market clearing – supply S is exogenous and must meet current demand. Higher demand drives up returns. Returns are, however, negatively related to past demand ϕ_{t-1} through the expected value of an asset – a *valuation mechanism*. If there is a positive social signal, investors still value an asset highly, even in the presence of low returns.

To explain the basic intuition, we consider the following scenarios: i) one where the asset has a positive return r_t and no social signal ϕ_{t-1}, and ii) one where investors observe a positive social signal ϕ_{t-1}. In scenario (i), demand is driven by the extrapolation component alone – investors believe that returns will continue to increase based on the current trend. In scenario (ii) on the other hand, investors do not require a large return to demand the asset – positive past sentiment drives current demand ϕ_t. Under exogenous supply, a strong positive signal from peers means that the extrapolated return is *less* important in justifying a higher price. The underlying reasoning relies on the fact that the system is in equilibrium.

Therefore, both returns and sentiments have adjusted to reflect a new steady-state, where sentiments are at a certain level ϕ_t.

Finally, we observe that the ratios of the coefficients, γ/β and α/β, play an important role. β effectively anchors the demand of investors in reality – a greater value of β implies that social signals carry less weight, and investors focus on price trends to forecast and expect asset values to grow at some constant rate. As β decreases, returns are determined more by social forces – hype from peers, rather than past performance, now justifies returns and demand. In our data, we observe that β is roughly five times α – individuals weight the sentiments of peers, however, returns are necessary to justify their investment strategies. σ^2 serves to taper the impact of sentiment, since investors are less certain in their signal and demand less of the asset.

Persistent fluctuations The reversal in returns is an important feature that emerges from social contagion in investors' price expectations. If large enough, these can produce bubbles in asset prices: initial momentum from positive news creates a price run-up, before an absence of news creates a drought of new asset demand. The subsequent price crash carries on its own momentum. We can treat demand as a latent variable to see these oscillations manifest in return data. Substituting lagged demand into the equation for returns, and iterating infinitely yields

$$r_t = -\sum_{T=1}^{\infty}\left(\frac{\alpha}{\gamma\sigma^2}\right)^T r_{t-T} + \frac{S_t\gamma\sigma^2}{\beta N} \tag{22}$$

as long as $\alpha/\gamma\sigma^2 < 1$, so that the contribution of demand fluctuations to returns converges to zero over time. This is an autoregressive model with infinite lags, where the coefficients decrease exponentially with lag size T. Without any knowledge of asset demand, the second term encapsulates an unobservable error term, which the model links to exogenous changes in the asset's supply. Eq. 22 demonstrates that an exogenous increase in returns at time t is followed by a smaller decrease in $t+1$. This oscillation persists indefinitely, and would converge to zero rapidly if the social signal $\alpha/\gamma\sigma^2$ is sufficiently small.

3.4 Model predictions

A simple linear regression exercise in Table 2 provides some evidence of the validity of the model proposed by Eq. 21. We summarise our asset demand model with social contagion by three further predictions, which we seek to validate in our WSB data.

Prediction 1: A mechanism for peer effects in asset demand *Given that asset demands by social investors are complementary, expressed sentiments react accordingly. A marginal increase(decrease) in peer outlook about an asset will raise(lower) the future outlook of an investor about the asset.*

We dedicate Section 4 to investigating strategic complementarities among investors on WSB. Besides testing for the direct effect of peers on investor sentiment (Prediction 1), we also use the opportunity to test our assumption for mechanical extrapolation. A uniform, marginal increase(decrease) in an asset's returns will raise (lower) the future outlook of an investor about the asset. It will also indirectly increase(decrease) the outlook of an investor through increasing(decreasing) the outlook of his peers.

Prediction 2: Return predictability and reversals *Stock returns are higher(lower) when asset demand is large(small) within a time period.*

Eq. 21 predicts that returns and WSB sentiments correlate positively contemporaneously, but negatively with respect to lagged sentiments. One issue, with regards to the positive correlation, is identification: the data will likely reflect the equilibrium outcome, in that sentiments are positive because returns are positive, and vice-versa. One challenge is to find variations in current sentiment that are exogenous with respect to current returns. Stock-specific characteristics will also drive persistent heterogeneity in the expressed sentiments of WSB users. We tackle the question of whether social dynamics can account for some return predictability in Section 5.

Prediction 3: Granular social forces *Idiosyncratic social shocks do not average out, and impact asset prices.*

Eq. 15 predicts that, in the presence of heavy-tailed attention to certain investors, heterogeneity in investor sentiments will not average out, and will instead have an impact on asset prices. An emphasis on viral content compels WSB discussants to follow specific, popular strategies, which are predicted to have an outsized impact on asset price returns. We investigate the role of these granular idiosyncratic sentiments in Section 5.

4 Social dynamics in WSB

This section provides empirical evidence for the existence of two mechanisms underlying asset demand – namely peer effects and extrapolation – among investors on WSB. We convey our main intuition using a game with strategic complementarities in retail investor decision-making (Zenou 2016, Hellwig & Veldkamp 2009, Bulow et al. 1985). Barlevy & Veronesi (2000) argue, contrary to Grossman & Stiglitz (1980), that learning among investors can become a strategic complement. Section 3 proposes the framework. We seek to test whether these complementarities manifest in the sentiments expressed about the future outlook of an asset among investors on WSB.

Testable prediction Prediction 1 in Section 3 establishes the behaviours we expect to see within the WSB community. In this section, we argue that user sentiment data observed on WSB are consistent with our model: investors are influenced by peer sentiments, and extrapolate past returns. WSB, as a platform, is a venue for 'social investors' to realise their strategic information complementarities.

Estimating equation The target independent variable of interest for studying hype investor sentiment is the log-odds of bullish over bearish sentiment,

$$\Phi_{i,t} = g(b_{i,t}) + f(\bar{\phi}_{-i,(t-1,t)}) + \varepsilon_{i,t}, \tag{23}$$

derived from our utility framework in Appendix B.1. One key addition is the time subscript, t. An author chooses a bullish over bearish strategy depending on: i) a signal $b_{i,t}$, and ii) the observed sentiments of peers, $\bar{\phi}_{-i,(t-1,t)}$.

4.1 Empirical strategy: consensus formation among investors

We use two approaches to estimate Eq. 23: i) the *Frequent Posters* approach, and ii) the *Commenter Network* approach. Both leverage different features of our data. For the *Frequent Posters* approach, we leverage the fact that certain users post multiple submissions about the same asset (hence, *frequent*). For the *Commenter Network* approach, we use instances in which users comment on others' submissions to more precisely gauge the transmission of sentiments about the same asset.

For the *Frequent Posters* approach, we observe that 8,173 authors create at least two submissions about the same ticker. We quantify peer influence by identifying the impact of other authors who write submissions about the same asset *between* an individual's two submissions. We use an IV of previous, expressed peer sentiments to control for exogenous shocks (see Figure 2 for an illustration). Our approach allows us to control for the author's sentiment prior to exposure to his peers, in addition to market moves.

We argue that the peer sentiments that an individual is exposed to have random, temporal variation: the posts that an individual is exposed to on WSB depend on what other anonymous, disconnected users have posted on the forum shortly before the author logs on, and what topic has recently gained popularity (see Section 2 and Appendix A.1 for a detailed description). Users are 'disconnected' in the sense that Reddit does not have friendship/follower ties. Individuals cannot, therefore, filter exposure to certain sentiments over others. We argue that this randomised exposure of users to different opinions (similar in spirit to random assignment of individuals to groups, such as in Weidmann & Deming (2021)) allows us to estimate direct peer effects.

The *Commenter Network* approach considers a submission-to-submission network, with an earlier submission exerting peer influence on a future submission if the author of the later submission commented on the earlier one. The submission-to-submission network helps identify peers an author interacts with more precisely. Here, we also control for market moves, and employ a set of IVs to address endogeneity concerns. As our IVs, we measure: i) sentiments of submissions to which the influencing submission is connected (the 'friends of friends' – detailed in Figure 3b), and ii) the historic sentiment of neighbours. The underlying argument rests on the premise that neighbours of network distance two exert an influence on user sentiments through peer effects (consistently with Bond et al. (2012)). A user's endogenous choice to comment on certain posts over others would therefore not account for users one step removed.

4.1.1 Identifying peer influence: Frequent Posters

Within WSB, we observe author i initially express a sentiment about an asset j, $\Phi_{i,j,(t-1)}$ (the continuous log-odds of a post expressing bullish over bearish sentiment, as per Eqs. 1&23), and, subsequently, write a new submission about the same asset at a later time, with an updated sentiment $\Phi_{i,j,t}$ (where time t is in event time). In the time between these posts, the author may observe submissions by others on the same asset expressing average sentiment $\bar{\Phi}_{-i,j,(t-1,t)}$, in addition to outside information related to the asset. Our goal is to identify the effect that expressed peer sentiments have on changing author i's sentiment.

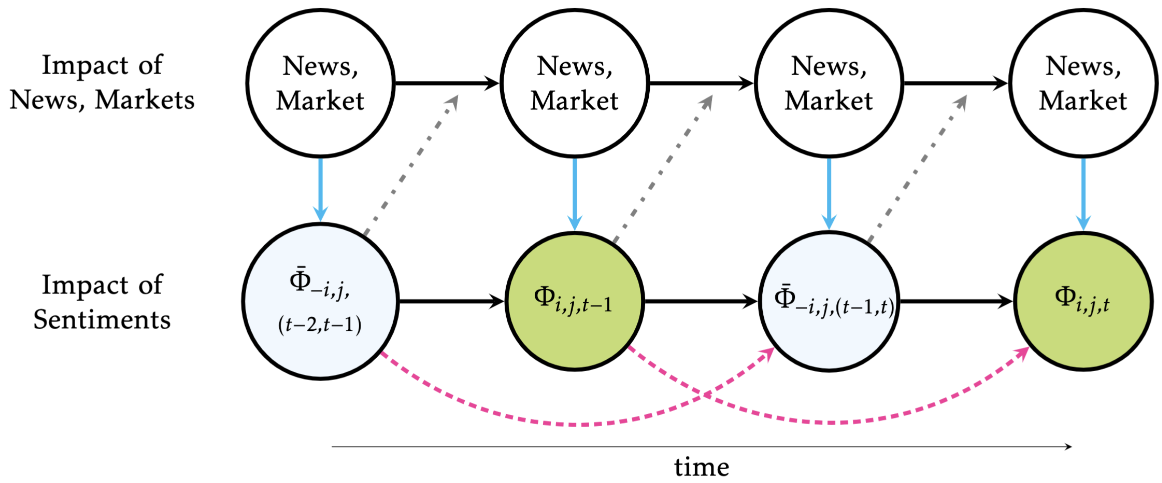


Figure 2: ***Frequent Posters*** **Directed Acyclic Graph (DAG);** we trace the flow of information within our system. Arrows represent the impact that information from one source has on the next source. Light blue nodes $\bar{\Phi}_{-i,j,(t-2,t-1)}$ and $\bar{\Phi}_{-i,j,(t-1,t)}$ represent peer sentiment; green nodes $\Phi_{i,j,t-1}$ and $\Phi_{i,j,t}$ represent the sentiment of investor i - our target variable. Time t is expressed in event time. The magenta, dashed line represents the impact that historical sentiment expressed about an asset has on the author's own future opinion. In our first stage, we estimate $\bar{\Phi}_{-i,j,(t-1,t)}$ by $\bar{\Phi}_{-i,j,(t-2,t-1)}$, controlling for the market move at the time that of the peer's initial post while estimating the coefficients. In this way, we are able to isolate the impact that peer sentiment $\bar{\Phi}_{-i,j,(t-1,t)}$ has on individual i at time t, $\Phi_{i,j,t}$.

Reduced form We first estimate the effect of average peer sentiment between an author's two submissions with the following linear model:

$$\Phi_{i,j,t} = \kappa\bar{\Phi}_{-i,j,(t-1,t)} + X_{i,j,t}\beta + \epsilon_{i,j,t}, \tag{24}$$

where the vector of control variables, $X_{i,j,t}$, is composed of stock-specific fixed effects, author i's past sentiment, and stock log-returns, both on day t and the average of the five days preceding t, and the variance in log returns on the five days prior to day t; β is a vector of corresponding coefficients. Even though peers appear randomly on the forum in this formulation (as discussed earlier in this section), an exogenous shock in the period $(t-1,t)$ may affect the views of both peers and the author in question simultaneously. For this reason, the OLS estimates do not enable us to precisely estimate peer influence.

Instrumenting peer sentiment I To tackle this issue, we use the historical views of peers as an IV for their views expressed within $(t-1,t)$. Our choice of IV is founded in psychology: Ross et al. (1975) find that 'once formed, impressions are remarkably persevering and unresponsive to new input', with later studies, such as Anderson et al. (1980), supporting these findings. We reason about our choice of IV through the Directed Acyclic Graph (DAG) shown in Figure 2. We consider that historic news and market moves are fully reflected in the news and market information available within the following timestep. Information shared by peers is also fully incorporated from one timestep to the next; however, dotted pink lines indicate the persistence of individual author sentiments (the persistence of individual formed impressions).

Leveraging the structure of our DAG, we estimate investor k's sentiment (a peer of investor i) about asset j, $\Phi_{k,j,t}$, based on the sentiment they expressed previously, $\Phi_{k,j,t-1}$, and control for asset returns at the time of their original post, $r_{j,t-1}$:

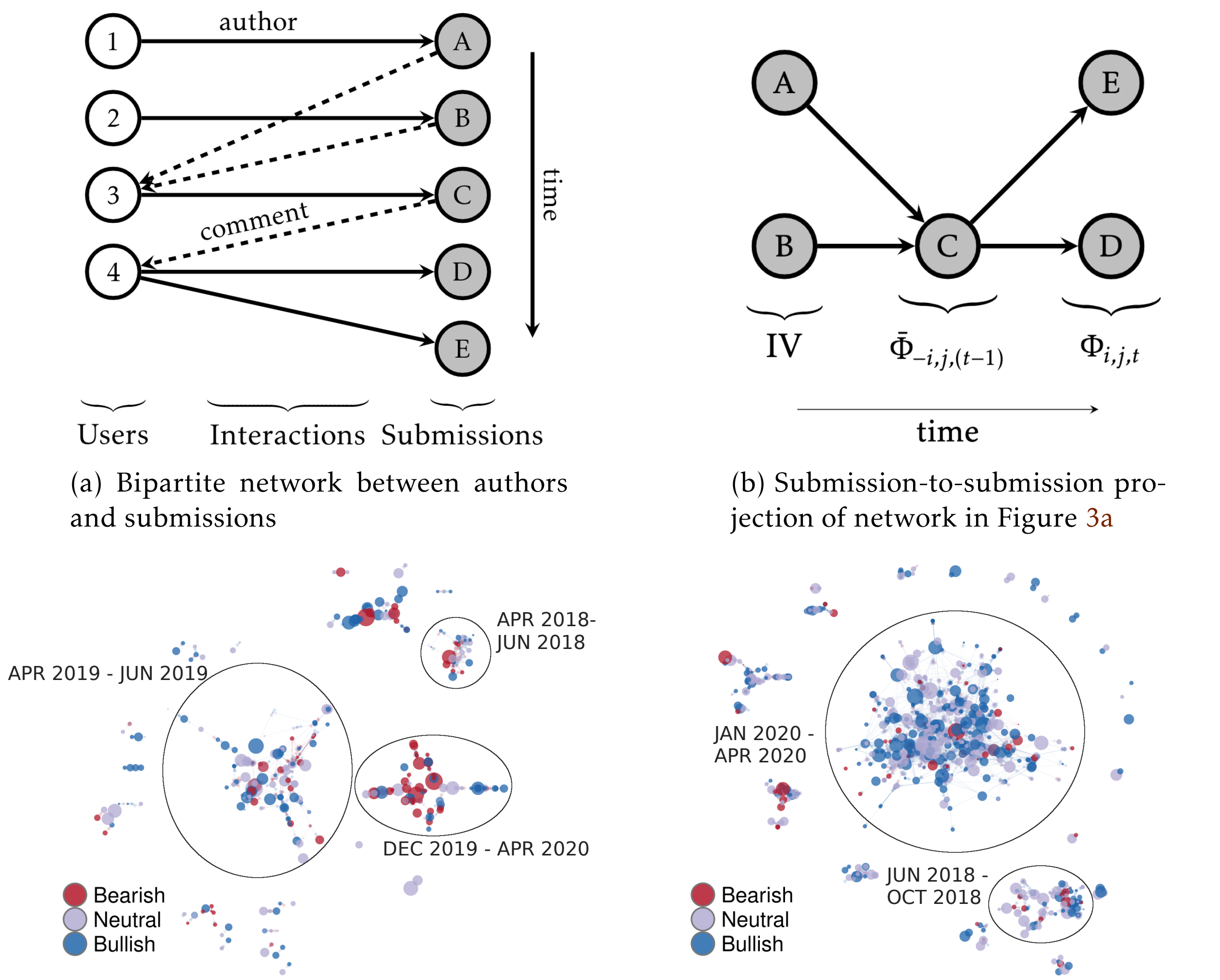
$$\Phi_{k,j,t} = \kappa_1^0 \Phi_{k,j,t-1} + \kappa_2^0 r_{j,t-1} + \epsilon_{k,j,t}^0, \tag{25}$$

where $\epsilon_{k,j,t}^0$ is an idiosyncratic error. The coefficient κ_1^0 estimates the true effect of an individual's historical sentiment. Consistently with our DAG, controlling for $r_{j,t-1}$ allows us to accurately estimate κ_1^0, while controlling for confounders. Eq. 25 is estimated using a sample containing submissions by all authors who post multiple times. The F-statistic for this first stage estimate, presented in Panel B.1 of Table 3, suggests that this is a strong instrument. Our choice of IV gives a good approximation for author sentiment, while allowing us to control for common shocks affecting the sentiments of peers and investor i in the period $(t-1, t)$. We use the predicted outlook of peers between an author's posts, $\hat{\Phi}_{-i,j,(t-1,t)}$, to estimate peer effects as our Second Stage regression, while keeping all other controls the same – historic peer sentiment is used for prediction. Appendix B provides further details on our variable construction and method. Appendix A.4 describes the construction of market variables, and their matching to WSB data.

Credible estimation We check whether our estimation strategy is credible, with respect to the three challenges presented by Zenou (2016) and Athey & Imbens (2017) in estimating peer effects. The first lies is in distinguishing peer effects from contextual effects – the tendency of perspectives to vary with some observable characteristics of the group, rather than individuals influencing each other. Controls for asset price movements and ticker specific characteristics – the main sources of exogenous variation – address this concern. Second, the random, anonymous nature of WSB, as well as controlling for ticker-specific fixed effects, address the possibility for correlated effects. The specification with the IV addresses the common shock problem. A more rigorous, statistical analysis of our identification strategy is included with the results.

4.1.2 Identifying peer influence – Commenter Network

WSB allows us to trace the interactions of users through a commenting network, even though there are no user friendship ties. We exploit a submission-to-submission interaction network for each asset, tracking which submissions in the past influence future submissions based on authors' commenting histories. This method offers a more precise way to identify a user's peers by observing which individuals, and submissions, an author explicitly interacts with. Figures 3a and 3b illustrate the approach.



(a) Bipartite network between authors and submissions

(b) Submission-to-submission projection of network in Figure 3a

(c) Submission-to-submission network: DIS

(d) Submission-to-submission network: MSFT

Figure 3: **User networks in WSB conversations**; WSB data is summarised as a bipartite graph, illustrated in Figure 3a, where users (left) are linked to submissions (right) when they author the submission (solid edge) or comment on the submission (dashed edge). The resulting projection of submissions, in Figure 3b, tracks the propagation of sentiments Φ. The submission-to-submission networks for two stocks in Figures 3c and 3d reveal that individuals post more submissions that are bullish(bearish) at times when the price of an asset increases(decreases) dramatically, with some visual evidence that similar sentiments tend to cluster.

Two examples of submission-to-submission networks in our data are displayed in Figures 3c and 3d. Distinct temporal clusters emerge, as a certain asset gains and loses prominence on WSB. Some discussions appear fragmented: the *DIS* discussion in Figure 3c, for example, contains several smaller clusters, with perceptible differences in overall sentiments. Others, such as the *MSFT* discussion in Figure 3d, contain a giant component where investors with different sentiments interact.

Our network approach uses a similar Reduced Form and Second Stage to the *Frequent Posters* approach in Eq. 24. We modify our control for an author's past sentiment about the stock to account for authors who post for the first time: a dummy variable encodes whether the author's most recent previous post is bearish, neutral, bullish or missing.

Instrumenting peer sentiment II We use an IV approach to estimate peer influence. As the First Stage, we estimate the sentiments of neighbours to estimate an author's view. As indicated in Figure 3b, the sentiments in submissions **A**, **B** can be used to predict that of submission **C**. The *predicted* sentiment of **C** can then, in turn, be used to predict the sentiments of **D** and **E**. This choice of IV is well-established in the networks literature (Zenou 2016, Patacchini & Zenou 2016, Bifulco et al. 2011), and helps control for the exogenous choice to comment on certain submissions and not others. We also include the neighbour's own historical sentiment, as a set of categorical variables, as the second IV (similarly to the *Frequent Posters* approach). Our Eq. 25, therefore includes a set of author controls $X^0_{i,j,t}$:

$$\Phi_{k,j,t} = \kappa^0 \Phi_{k,j,t-1} + X^0_{i,j,t}\beta^0 + \epsilon^0_{k,j,t},$$

where the superscript denotes the estimation of the First Stage. In the results in Table 3, we display the estimate for our main IV - neighbours of neighbours in the commenting network; the additional IV of author historical sentiment is displayed in Appendix B.

Timing of observations We use the timings of events to mitigate the common shock problem for both our IVs: the neighbour's historical sentiment and the 'friends or friends' submissions. For the latter, we calculate the time period of influence for a given post, which ends when the last comment is made on a submission. This effectively marks the point when a particular submission ceases to be of interest to the WSB community. We filter for instances where the period of influence for a submission used as an IV for another submission ends before the new submission we are modeling is created. In practice, if submission **C** in Figure 3b occurs on July 1st at 2:31PM, the final comments on posts **A** and **B** must occur before, in order to ensure that our IV is not affected by a common shock. We also include an author's own, historical sentiment as an IV only if his previous submission occurs at least two business days before the current one.

The *Commenter Network* offers certain upsides, but also certain shortcomings, as compared to the *Frequent Posters* approach. The network method more precisely identifies the channels of influence between authors. However, the allocation of peers is no longer random, since the network structure is governed by a *choice* to comment.

4.2 Results: peer effects and extrapolation

In this section, we present the Reduced Form, Second Stage, and First Stage regression estimates for both the *Frequent Posters* and *Commenter Network* approaches. The Reduced Form and Second Stage estimates, across both model specifications, show that peer sentiments directly impact an individual's sentiment about an asset, with user sentiments conforming to those of their peers.

Table 3 presents the results, with Panel A presenting OLS estimates for κ, from Eq. 24, using observed variation in peer sentiments, and Panel B.1 using predicted variation in peer sentiments. We relegate estimated coefficients for control variables to Appendix B. In the

Table 3: Peer influence in WSB sentiments

	Frequent Posters (1)	Network (2)
Panel A: Reduced Form – peer influence estimated using *observed* average sentiment of peers		
	Dependent Variable: Investor Sentiment ($\Phi_{i,j,t}$)	
Average peer sentiment, $\bar{\Phi}_{-i,j,(t-1)}$ *(observed)*	0.10 (0.02) ***	0.05 (0.01) ***
$r_{j,t}$	0.85 (0.13) ***	0.84 (0.12) **
Author & asset controls ($X_{i,j,t}$)	Yes	Yes
Number of obs.	14,396	24,963
F-statistic	71	668
Panel B.1: Second Stage – peer influence estimated using *predicted* average sentiment of peers		
	Dependent Variable: Investor Sentiment ($\Phi_{i,j,t}$)	
Average peer sentiment, $\hat{\Phi}_{-i,j,(t-1)}$ *(predicted)*	0.19 (0.05) ***	0.19 (0.08) **
$r_{j,t}$	0.96 (0.17) ***	0.98 (0.24) ***
Author & asset controls ($X_{i,j,t}$)	Yes	Yes
Number of obs.	11,122	16,521
J-statistic	NA	0.43
F-statistic	77	1,277
Panel B.2: First Stage – estimating peers' sentiments		
	Dependent Variable: Sentiment of Peers	
Historical Sentiment of Peers	0.32 (0.01) ***	
Sentiment of Neighbours' Neighbours		0.14 (0.01) ***
Author controls ($X^0_{i,j,t}$)	No	Yes
Controls for returns ($r_{j,t}$)	Yes	No
Number of obs.	19,748	24,013
F-statistic	1,225	118

Notes: this table presents the First Stage, Second Stage and Reduced Form OLS estimates for peer influence on WSB. In column (1), the First Stage is estimated using the initial sentiment expressed by an author about an asset to estimate his sentiment in the following post. In column (2), the First Stage is estimated using the sentiment of previous submissions that an author commented on, regarding the same asset. The Second Stage is estimated using the average predicted sentiment of peers. Ticker-level dummies, asset return and volatility controls, and the intercept are included in the Second Stage and Reduced Form estimates, but not shown here; additional author-specific IVs in the network approach are also included but not shown – the complete estimates are presented in Appendix B. Robust standard errors, clustered at the ticker level for Panels A and B.1, are presented in parentheses. Observations with incomplete data are dropped.
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Frequent Posters approach, the estimated peer effects can be summarised as follows: an average estimate for κ at 0.1, in the Reduced Form case, means that doubling in the odds of peers expressing bullish over bearish sentiments increases the odds of a given submission to be bullish, over bearish, by 7.2%, on average (we raise the log-odds estimate for $\Phi_{i,j,t}$ to an exponent). In the IV setting, an estimate of 0.19 translates to an increase in the corresponding odds for bullish over bearish sentiments by 14.1%. In all cases, the robust standard errors, clustered at the ticker level, produce estimates statistically significant at the 1% level. The

Commenter Network approach yields a similar result.

The estimated coefficients in columns (1) and (2) of Panel B.1 suggest that an exogenous increase in average peer outlook appears to increase an investor's own future view about an asset. These findings demonstrate that the data are consistent with Prediction 1. As a result, we conclude that the data supports a model where strategic complementarities govern the investment decisions of retail traders sampled on WSB. The results in Table 3 also support our model with extrapolation. We observe that recent returns are highly predictive of expressed sentiments across all specifications.

Support for identification One potential concern is that individuals who post multiple times about the same asset, or those who comment on others' submissions, may differ from the rest of the population on the forum. If this were the case, our findings would not allow us to draw valid conclusions about the overall population of investors. We provide evidence that sentiments expressed by our samples are similarly distributed to those of the overall user population in Appendix B.

A second concern is whether our proposed independent variables – asset price movements, ticker fixed effects and author historical sentiments – are effective controls for unobserved ticker characteristics. If our controls in the *Frequent Posters* formulation are valid, then a randomly selected cohort of individuals who post on the same ticker *before* the author's first post, should have no effect on the sentiments expressed in dependent submissions. Similarly, if our controls are useful in the *Commenter Network* formulation, a random rewiring of the network should yield no effect. The results are detailed in Appendix B: no statistically significant correlation emerges from the randomly selected cohorts. This provides further evidence that unobserved factors influencing within-ticker variation in both peer composition and author sentiment are not confounding.

We also observe that our instrumented independent variable (Panel B.1) appears to be more significant than our uninstrumented one (Panel A). The data shows that *observed* peer sentiments $\bar{\Phi}_{-i,j,(t-1)}$ is correlated with $r_{j,t}$. This is not surprising, given the fact that we observe mechanical extrapolation. For this reason, in Panel A, the coefficients on both returns and peer effects are lower. Panel B.1 allows us to disentangle peer effects and recent market returns. Therefore, we observe a higher coefficient on both returns and *predicted* peer sentiment.

A final concern with our *Commenter Network* approach is overidentifying restrictions. A J-statistic of 0.43, and a corresponding p-value of 51%, leads us to believe that our additional instruments are exogenous (see Appendix B for further details). We explore further dynamics observed on WSB, such as whether there is contagion in asset interest among investors online (Banerjee 1993, Shiller 2017), in Appendix B.

4.3 Further insights

WSB data provide additional opportunities to test investor responses to a market surprise, and the reinforcement mechanism between peers and asset prices. We consider the sentiments expressed by investors i about asset j at time t, $\Phi_{i,j,t}$, as our dependent variable and use the controls from Eq. 24 to test for two additional effects.

We define two types of surprises: i) a positive surprise if asset j experiences a return which is two standard deviations higher than the 30-day historical average for the stock on day t or on the day before, and ii) a negative surprise if asset j experiences a return which is two standard deviations lower than the average for the stock on day t or on the day before. We compute the average and standard deviation for stock j using data of the thirty trading days before t. We also interact returns and predicted peer sentiment to see the extent to which peer effects are reinforced by returns. We use the *predicted* peer sentiment from our *Frequent Posters* approach in our regressions to control for sentiments that respond to current price changes.

Table 4: Additional effects: surprise and reinforcement

		Dependent Variable: $\Phi_{i,j,t}$	
		(1)	(2)
Independent Variables	$\hat{\Phi}_{-i,j,(t-1,t)}$	0.19 (0.05) ***	0.19 (0.05) ***
	$r_{j,t}$	0.85 (0.20) ***	-0.13 (0.39)
	Positive Surprise	-0.11 (0.06) *	
	Negative Surprise	-0.15 (0.05) ***	
	$(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^{+}$		1.49 (0.52) ***
	$(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^{-}$		-6.84 (5.45)
	Author & asset controls ($X_{i,j,t}$)	Yes	Yes
	No. Observations:	11,081	11,081
	R^2:	0.08	0.08
	R^2_{adj}:	0.06	0.06

Notes: The dependent variable – individual investor sentiment about an asset, scaled continuously between $(-\infty, \infty)$ – is estimated using the variables in Eq. 24 and additional variables, using OLS. The additional variables in column (1) are categorical variables for positive and negative market surprises at time t in asset j; in column (2) the additional variables are a cross term between asset j's returns and the estimated sentiments of peers: $(r_{j,t}\ \mathrm{X}\ \hat{\Phi}_{-i,j,(t-1,t)})^{+}$ is the product if the predicted sentiment of peers is positive and zero otherwise, and captures the extent to which positive peer predictions correspond to observed market moves; the reverse is true for $(r_{j,t}\ \mathrm{X}\ \hat{\Phi}_{-i,j,(t-1,t)})^{-}$. Peer sentiment $\hat{\Phi}_{-i,j,(t-1,t)}$ is estimated using the *Frequent Posters* approach to control for confounders. Robust standard errors, clustered at the ticker level, are presented in parentheses. Observations with incomplete data are dropped.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Surprises Table 4 presents the results from our exploration of surprise and reinforcement. Column (1) contains the OLS estimates when including positive and negative categorical variables for market surprise. A negative market surprise appears to significantly affect investor sentiments. The result is not symmetric – a positive surprise does not appear to convince investors of the upside potential of a stock, but rather has a negative effect on sentiments (the confidence in this estimate, however, is fairly low). This observation suggests that downside panic spreads quickly within the investor population. This effect is in addition to the large impact returns have on sentiment.

Reinforcement Column (2) considers the effect from market reinforcement of peer sentiments by including the cross term between returns and the predicted sentiments of peers. The cross terms are separated depending on whether the predicted peer sentiment $\hat{\Phi}_{-i,j,(t-1,t)}$ is positive or negative: $(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^+$ is the *bullish* interaction when $\hat{\Phi}_{-i,j,(t-1,t)}$ is positive and zero otherwise, whereas the bearish interaction $(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^-$ is zero if predicted sentiment takes a negative value. Therefore, a large value for the bullish interaction corresponds to peers forecasting positive returns in asset j and the asset j simultaneously experiencing positive returns on the day of author i's submission.

In Table 4, the bullish interaction is highly significant. WSB users are spurred by peers predicting positive returns and subsequently observing the asset outperform in the market, possibly suggesting some 'irrational exuberance' (Shiller 2005). The reverse is not true for bearish reinforcement. Even though the coefficient for interaction $(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^-$ is negative, it is not statistically significant. Users correctly forecasting downturns is not found to be convincing for the typical WSB discussant. Perhaps equally important is the coefficient on returns $r_{j,t}$, which does not add any statistical power in this regression. This underscores the important role that peers play in spreading information about assets, and the fact that asset performance alone does not garner enthusiasm.

Summary There are several implications to consider when looking at the findings in Columns (1) and (2) in Table 4 simultaneously. The spread of upside enthusiasm and fear of downside exhibit fundamentally different properties. Peer excitement about an asset's potential is a key mechanism for spurring investor enthusiasm. However, information about downside risk appears to have relatively little effect. An asset experiencing large losses, however, triggers fear and pessimism among retail investors on WSB.

5 Has WSB destabilised markets?

Prediction 2 states that strategic complementarities impact asset returns. In Section 4, we focus on identifying the extent to which investors consider sentiments from peers, as well as recent market moves, while updating their own sentiments about future asset returns. The evidence for peer effects in asset demand is robust in two, separate estimation strategies we consider. In this section, we conduct a quantitative analysis of asset returns and the data from WSB, in order to validate the link between these behaviours and stock returns.

The key challenge is that sentiments and returns are co-determined in equilibrium. If returns are high, sentiments are also high. Conversely, if sentiments are high, buying pressure will also increase returns contemporaneously. We consider two empirical strategies to investigate the extent to which WSB *caused* returns to exceed a benchmark without social contagion. The first strategy exploits variation in sentiments that we can explain using the history of WSB conversations. The second strategy exploits the granularity of discussions on WSB.

5.1 Return predictability through persistent demand

We are interested in finding variation in current sentiments which are exogenous with respect to current returns. The goal is to detect a positive effect carried by retail investor sentiment, proxied by WSB activity. First, we formulate the linear relationship between returns and current sentiments. We then propose a 2SLS estimation strategy to quantify the impact of social contagion on stock market variables.

Independent variable Our independent variable estimated from WSB sentiment data measures sentiment change $\Delta\Phi_{j,t}$, the first-difference of stock j's mean daily sentiment between calendar weeks t and $t-1$. The purpose for this variable is to gauge the stock-specific response to a change in WSB's associated attention and sentiments on a week-by-week basis. Measuring the difference in average sentiments between two periods proxies for the change in asset demand due to changes in the intensity of corresponding sentiments.

Reduced Form We regress changes in weekly log-returns on changes in weekly sentiments:

$$\Delta\bar{r}_{j,t} = \omega\Delta\Phi_{j,t} + \eta_t + \varepsilon_{j,t}, \tag{26}$$

where ω is the coefficient of interest, η_t denotes week fixed effects, and $\varepsilon_{j,t}$ an idiosyncratic error. This Reduced Form setup does not allow us to argue that a causal relationship exists between social investor activity and stock market activity: the narratives discussed and sentiments expressed at a given point in time are often shaped by real-time news, events and stock market moves, resulting in reverse causality. For example, users express positive sentiments in weeks of outsized, positive returns, regardless of past sentiment.

First Stage We use variation in sentiments that can be explained by past activity on WSB and past stock performance to identify our parameter of interest. In doing so, we assume that log returns are sufficiently uncorrelated between sequential trading weeks. Effectively, this translates to an assumption that the market is sufficiently efficient, such that stocks' weekly returns are uncorrelated to those in the previous week.

We predict sentiment $\Phi_{j,t}$ using past stock price behaviour, as well as past sentiments:

$$\Phi^+_{j,t} = \log\left(\frac{\mathrm{P}(\phi_{j,t} = +1)}{\mathrm{P}(\phi_{j,t} = 0)}\right) = \lambda^+_r \bar{r}_{j,t-1} + \lambda^+_\sigma \sigma^2_{j,t-1} + \lambda^+_1 \Phi^+_{j,t-1} + \lambda^+_2 \Phi^-_{j,t-1} + \eta^+_t + \varepsilon^+_{j,t}, \tag{27}$$

$$\Phi^-_{j,t} = \log\left(\frac{\mathrm{P}(\phi_{j,t} = -1)}{\mathrm{P}(\phi_{j,t} = 0)}\right) = \lambda^-_r \bar{r}_{j,t-1} + \lambda^-_\sigma \sigma^2_{j,t-1} + \lambda^-_1 \Phi^+_{j,t-1} + \lambda^-_2 \Phi^-_{j,t-1} + \eta^-_t + \varepsilon^-_{j,t}, \tag{28}$$

where superscripts differentiate between the average log-odds of a submission in week t expressing bullish (+) versus negative (−) sentiments, over neutral sentiments. Week fixed effects remain in the sentiment models, so that the full estimation strategy rests on within-week variation in all explaining, as well as explained, variables.

The approach outlined above relies on coarse aggregates for sentiments: the probabilities here are not estimated on data for individual submission sentiments, as is the case in Section

4. Rather, the probabilities are calculated by averaging the probabilities for *all* submissions in week t, discussing ticker j, to be bullish ($P(\phi_{j,t} = +1)$), bearish ($P(\phi_{j,t} = -1)$), or neutral ($P(\phi_{j,t} = 0)$). Predicted values for our sentiment measure follow from the fitted sentiment model:

$$\Delta\widehat{\Phi}_{j,t} = \frac{1}{2}\left(\widehat{\Phi}^{+}_{j,t} - \widehat{\Phi}^{-}_{j,t}\right) - \Phi_{j,t-1}, \tag{29}$$

where a hat denotes the values fitted from the first stage regressions.

Results In all our estimates, we restrict ourselves to a sub-sample spanning January 2016 to July 2020. This choice serves to limit the amount of missing data in times when activity on WSB was relatively sparse.

Table 5 helps assess the instruments' strength in predicting sentiments on WSB. The high F-statistics justify that the explanatory variables are not weak instruments. In both columns, we find that lagged weekly mean and variance in returns, combined with lagged sentiments, are significant predictors for the current log-odds in average weekly submissions expressing bullish and bearish sentiments. This is in line with our findings in Section 4.

Table 5: First Stage estimates for consensus and contagion in WSB

	Dependent variable:	
	$\Phi^{+}_{j,t}$	$\Phi^{-}_{j,t}$
$\bar{r}_{j,t-1}$	−0.0002 (0.52)	-1.53^{**} (0.65)
$\sigma^{2}_{j,t-1}$	-3.78^{***} (0.98)	-3.75^{***} (0.68)
$\Phi^{+}_{j,t-1}$	0.09^{***} (0.02)	-0.06^{***} (0.01)
$\Phi^{-}_{j,t-1}$	-0.03^{***} (0.01)	0.16^{***} (0.01)
Week FE	Yes	Yes
Number of obs.	6,711	6,711
F-statistic	17.63	49.53

Notes: the dependent variable in Column (2) is the average log-odds of a given submission in week t on stock j to express bullish over neutral sentiment, and in Column (3) – bearish over neutral sentiments. Explanatory variables include: the average log-return $\bar{r}_{j,t-1}$, and the variance in log-returns $\sigma^{2}_{j,t-1}$. The logit-transformed sentiments are regressed on the lag of the weekly mean and variance of log-returns, as well as the lag in logit-transformed sentiments. Each specification includes week-specific fixed effects. Accompanying standard errors, displayed in brackets, are clustered at the stock level, and calculated in the manner of MacKinnon & White (1985).
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 6 presents our main results. Panel A regresses changes in average returns against *observed* measures for sentiment changes $\Delta\Phi_{j,t}$. Panel B in Table 6 presents causal evidence for the impact of sentiments among WSB users on stock market variables, using *predicted* sentiments from the model presented in Table 5. However, this average effect is small, which is not surprising given that many of the stocks discussed on WSB have large market capitalisation.

We do not argue that WSB alone affects the markets, but rather that WSB data offers a rich

Table 6: Market impact of WSB discourse

Panel A: Reduced Form relationship between WSB and market activity	
	Dependent variable:
	$\Delta\bar{r}_{j,t}$
$\Delta\Phi_{j,t}$	0.002*** (0.0003)
Week FE	Yes
Number of obs.	6,671
F-statistic	24.32
Panel B: structural relationship between WSB and market activity	
$\Delta\widehat{\Phi}_{j,t}$	0.004*** (0.001)
Week FE	Yes
Number of obs.	6,671
F-statistic	12.63
J-statistic	8.108

Notes: this table presents OLS estimates for stock j's change in average log-return, $\Delta\bar{r}_{j,t}$, in week t. We filter the sample to stocks mentioned in at least 31 distinct submissions on WSB, and exclude any ETFs. Explanatory variables include a measure for sentiment change, $\Delta\Phi j,t$, which tracks the change in average sentiments on WSB. Each specification includes week-specific fixed effects. Accompanying standard errors, displayed in brackets, are clustered at the stock level, and calculated in the manner of MacKinnon & White (1985). Panel A computes the coefficients using values directly from WSB data, whereas Panel B employs sentiments and stock discussion predicted by past sentiments, stock discussions, as well as returns and return volatility, for which results are in Table 5. The associated J-statistics are recorded at the bottom of Panel B, which are computed by regressing the residuals from the Second Stage on all variables used for predicted $\Delta\widehat{\Phi}_{j,t}$.
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

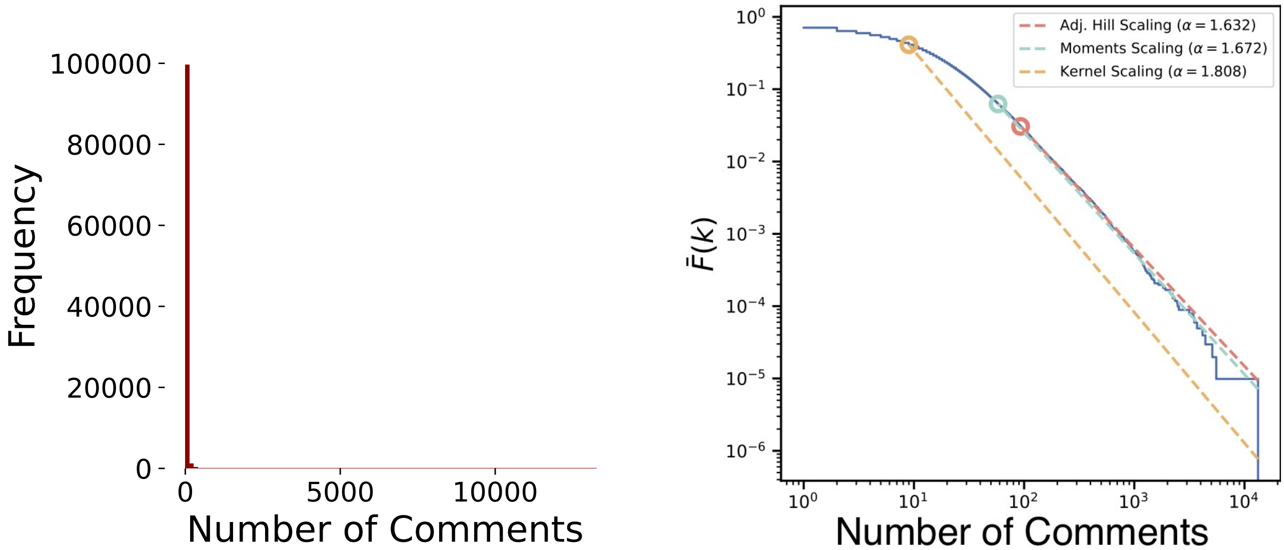
sample of retail investor behaviour. Variation in sentiments explained by the past offers a measure for the intensity by which retail investor asset demand propagates from one week to the next. Even though sentiments reflect current returns, prior beliefs are expected to change prices, thus returns, beyond the market average accounted for by time fixed effects.

5.2 Granular social forces

Thus far, our paper demonstrates that strategic information complementarities can drive persistence in sentiments and oscillations in returns. Our empirical exercise in Section 4 also shows that individual sentiments are not fully explained by information from peers or recent returns. There is, therefore, unexplained heterogeneity in individual investor sentiments. An outstanding question is whether these heterogeneous opinions can survive aggregation across peers and impact asset returns. In this section, we strive to test Proposition 3 from Section 3.

We leverage the heavy-tailed structure of WSB discussions for our empirical strategy. The intuition is that certain submissions gather many more followers than others, which we measure using the number of comments they receive. Figure 8 in Appendix A.2 displays the heavy-tail in discussions between assets, but sentiments *within* stocks also rely predominantly on a few, heavily-commented discussions. In our modification to our model, the heavy-tail of attention can result in unexplained variation in sentiment surviving aggregation and impacting returns. We use a Granular Instrumental Variable (GIV) approach to investigate the impact on returns (Gabaix & Koijen 2020, 2021, Galaasen et al. 2020).

Granularity of social attention Similarly to other online social phenomena, the attention granted to different submissions on WSB is heavy-tailed. A small fraction attract a high number of comments and upvotes, while most go virtually unnoticed.



(a) Distribution of comments across all submissions

(b) Tail estimate for the distribution of comments

Figure 4: **Attention is heavy-tailed**; These graphs show the distribution of comments that submissions that mention a ticker on WSB receive. The left figure plots the distribution. The right plots the tail exponent, estimated using the three different methods outlined in Voitalov et al. (2019). The tail exponent is estimated to be less than two across all methods – this implies that the tail distribution obeys a power law, and is heavy-tailed.

We begin by establishing that the distribution of attention that information shared by investors online receives is heavy-tailed. We proxy attention by the number of comments that a particular submission receives. Figure 4 shows the extreme tail in the distribution of attention – some submissions appear to receive a large following, while the majority are of little interest. We study the distribution of comments using the methods discussed in Voitalov et al. (2019), who propose several methods for estimating the power-exponent of a distribution's tail – all methods estimate the tail exponent to be less than two, implying that the tail is power-law and heavy-tailed. The heavy-tailed attention online implies that idiosyncratic information contained within the most popular submissions potentially persist after pooling across all investor's opinions, and may have a disproportionate effect on returns. For our identification strategy, we exploit within-ticker-week variation in attention. Therefore, we filter our sample to weeks and tickers where a sufficient number of submissions are made - we choose fifteen submissions for our cutoff.

Estimates of idiosyncratic social shocks The next step of our empirical approach consists in extracting idiosyncratic information shocks, measured as unexplained idiosyncratic variation in the sentiments of submissions. To extract unexplained variation in sentiments, we regress the sentiment expressed in a given post $\Phi_{i,j,t}$ on the return on day t, $r_{j,t}$, the average returns on the week containing t, r_{j,t_w}, and average sentiments expressed by peers in the prior week $\bar{\Phi}_{j,t_w-1}$.

The strategy follows the reasoning outlined in Section 4.2, Figure 2. We posit that any news that emerges at time t about a company is assimilated by the market and manifests in returns. Any variation in sentiment that is unexplained by market performance at time t and by past discussions is post-specific and is idiosyncratic to news and information more broadly available about that stock at that time. For a post about asset j made by author i at time t, we estimate the idiosyncratic information content of the post as $e_{i,j,t}$ in the following regression:

$$\Phi_{i,j,t} = \beta_1^A r_{j,t} + \beta_2^A r_{j,t_w} + \beta_3^A X_{t_w} + \beta_4^A \bar{\Phi}_{j,t_w-1} + e_{i,j,t}, \tag{30}$$

where X_j are asset fixed effects. We add asset fixed effects in order to control for the fact that certain sentiments about assets are persistent on WSB – for example, the forum's enthusiasm about TSLA. We note that the results remain similar with and without ticker fixed effects.

The object of interest, residual $e_{i,j,t}$, is information shared in the submissions that is orthogonal to asset j's returns at time t or within week t_w. We, therefore, would expect $e_{i,j,t}$ to have an impact on the market through social forces, rather than through purely informational content. Figure 5 plots the idiosyncratic social shocks. The distribution is somewhat asymmetric, and the modal, unexplained sentiment is bullish, but the left tail of discussions demonstrates the presence of intense bearish discourse.

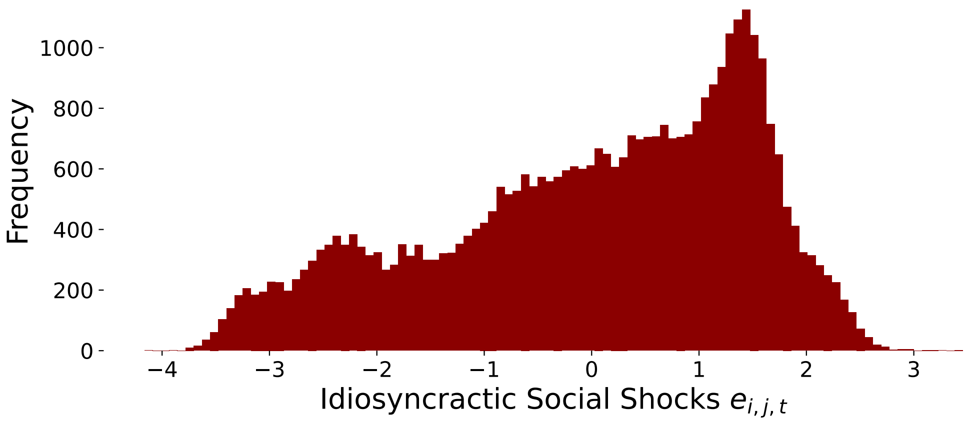


Figure 5: **Idiosyncratic social shocks:** this graph plots the distribution of idiosyncratic sentiment heterogeneity, $e_{i,j,t}$, from Eq. 30.

Estimating the effect of granular attention In order to assess the impact on asset returns of online attention, we proceed by analyzing the following relationship:

$$r_{j,t_w+1} = \beta_1^B \bar{e}_{j,t_w} + \beta_2^B \bar{\Phi}_{j,t_w-1} + \beta_3^B X_{t_w} + v_{j,t_w}, \tag{31}$$

where X_{t_w} are week fixed effects, r_{j,t_w+1} is the total return of stock j in week t_w, $\bar{\Phi}_{j,t_w-1}$ is the average sentiment expressed on the forum about asset j in week $t_w - 1$, $\bar{e}_{j,t_w} = \sum_i s^c_{i,j,t_w} e_{i,j,t_w}$ the popularity-weighted average idiosyncratic sentiment shared about stock j in week t_w ($\sum_i s^c_{i,j,t_w} = 1$), and v_{j,t_w} is a stock-week specific error. s^c_{i,j,t_w} is calculated by summing the total number of comments received across all posts in stock j in week t_w – the comment count for a specific submission is then normalized by the total comment count within week t_w; comment count is re-indexed so that the minimum number of comments a post receives is one.

A key identification challenge stems from our goal to identify the impact of social forces and popularity of some content over other content, versus general idiosyncratic sentiment on the forum. We rely on a Granular Instrumental Variable (GIV) for our identification strategy. The GIV is defined as the difference between popularity-weighted and equally-weighted social shocks, each aggregated for the stock at period t_w:

$$GIV_{i,t_w} = \sum_i^N s^c_{i,j,t_w} e_{i,j,t_w} - \sum_i^N \frac{1}{N_{j,t_w}} e_{i,j,t_w}, \tag{32}$$

where N_{j,t_w} is the total number of posts about stock j in week t_w. We subsequently replace $\bar{e}_{j,t_w}$ in Eq. 31 with $\hat{u}_{j,t_w}$, where $\hat{u}_{j,t_w}$ is the predicted values from the regression of the GIV on the social shocks $\bar{e}_{j,t_w}$. The outcome variable $\hat{u}_{j,t_w}$ is driven by the popularity of certain posts over others, rather than by general idiosyncratic sentiment. The use of the GIV also allows us to mitigate the common shocks problem, where certain stock-specific shocks could affect all idiosyncratic sentiments and future returns – this is discussed further in Appendix C.2, and is similar to Galaasen et al. (2020) in reasoning.

Results In order to study the financial consequences of granular social attention, we run the following regression on weekly returns and posts.

$$r_{j,t_w+1} = \beta_1^B \hat{u}_{j,t_w} + \beta_2^B \bar{\Phi}_{j,t_w-1} + \beta_3^B X_{t_w} + v_{j,t_w},$$

where $\hat{u}_{j,t_w}$ are the fitted values of idiosyncratic social shocks on our GIV, r_{j,t_w+1} is the cumulative weekly return of stock j in week t_w+1, X_{t_w} are week fixed effects, $\bar{\Phi}_{j,t_w-1}$ is the average sentiment expressed about stock j in the prior week. The formulation closely follows that of our extended model.

Several patterns emerge from our empirical exercised. Firstly, we observe that the data appears to follow the structure proposed in our model – idiosyncratic social shocks are positively linked to future returns, while historical sentiments are negatively linked. In the Column (5), the estimated effect can be summarised as follows: an estimate for β_1^B at 0.030, means that the idiosyncratic doubling in the odds of a very popular post expressing

Table 7: Stock returns versus granular social shocks

	Dependent variable: r_{j,t_w+1}						
	Average $\bar{e}_{j,t_w}$	Popularity-weighted $\bar{e}_{j,t_w}$	Instrumented by GIV: $\hat{u}_{j,t_w}$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Pooled	Pooled	Pooled	Pooled	Pooled	Positive	Negative
Granular Social Shock	0.022	0.026**	0.030**	0.031**	0.030**	-0.018	0.105***
	(0.016)	(0.011)	(0.013)	(0.013)	(0.013)	(0.027)	(0.034)
$\bar{\Phi}_{j,t_w-1}$	-0.001	-0.003		-0.012	-0.001	-0.007	-0.008
	(0.010)	(0.010)		(0.008)	(0.010)	(0.012)	(0.014)
Week FE	Yes	Yes	No	No	Yes	No	No
Controls in Eq. 31	Yes	Yes	No	Yes	Yes	Yes	Yes
Observations	916	916	916	916	916	229	229
R^2	0.244	0.248	0.006	0.008	0.247	0.004	0.042

Notes: this table presents the OLS estimates for the influence of idiosyncratic social shocks on WSB. Columns (1) and (2) present the effect of average idiosyncratic shocks and popularity-weighted idiosyncratic shocks, respectively. Columns (3), (4) and (5) present various specifications of our instrumented idiosyncratic social shocks $\hat{u}_{j,t_w}$. Columns (6), (7) estimate the effect of social shocks where the instrumented variable is in the top or bottom quartiles (Positive and Negative) of the estimated $\hat{u}_{j,t_w}$. Robust standard errors, clustered at the week level for columns (1), (2), (5), are shown in parentheses. The F-statistic for the first stage regression is 1239. Observations with incomplete data are dropped.
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

bullish over bearish sentiments (while less popular posts do not express an idiosyncratic sentiment) increases returns in the following week by one percent, on average. The effect is small, but persists across specifications. Consistently with our model prediction, the average idiosyncratic noise, in Column (1), has no effect on returns.

The effect of idiosyncratic social shocks is not symmetric, as seen in Columns (6), (7): returns respond to negative idiosyncratic social shocks, while positive ones have little effect. The effect of negative idiosyncratic shocks is approximately three times that of the average, and explains slightly over 4% of the return variation in the following week (as observed through the R^2). The disproportionate effect of downside social shocks reinforces our findings from Section 4.3 that downside panic about extreme losses spreads through the investor population and causes a shift in sentiments.

6 Conclusion

We contribute to the growing literature on asset demand models by documenting the influence of peers on retail investor opinion formation, using data from WSB as a case study. We specifically document complementarities in asset demand among retail investors. User sentiments are, on average, 14% more likely to be bullish rather than bearish, if the odds of peers expressing bullish over bearish sentiments double. These results are consistent with the findings of Pool et al. (2015) and Bursztyn et al. (2014).

We formalize the role that trend following and information from peers play in a model for asset returns. Our model highlights a diverse set of implications from these behaviours. In a multi-period equilibrium model for returns, we expect asset returns to oscillate in the short-term. Social investors drive up the price when they observe their peers investing greater amounts, since they gain greater confidence in their investment from peers.

In a quantitative exercise, we show that instrumented sentiments from WSB are closely linked to simultaneous returns, as predicted by the model. The instrument allows us to isolate the effect of WSB sentiments on returns, as opposed to the effect of returns on WSB sentiments. However, in the presence of 'granular' heavy-tailed attention, heterogeneous investment decisions are not averaged out and can impact returns. Using the method of Gabaix & Koijen (2020), we show that idiosyncratic sentiment heterogeneity among users (which is not reflective of fundamental news) impacts the market due the heavy-tailed nature of the popularity of online content.

Economists have long deliberated to what extent social dynamics and human psychology play a role in economic decision-making (Shiller 1984, Black 1986), with Hirshleifer (2001) concluding that 'despite many empirical studies, scholarly viewpoints on the rationality of asset pricing have not converged'. An outstanding question is, to what extent do social behaviours impact financial markets? We shed light on this question, using data from WSB as a case study for how retail investors behave, and hope to demonstrate concurrently the usefulness of online text data for understanding the social dynamics behind economic decisions.

References

Anderson, C. A., Lepper, M. R. & Ross, L. (1980), 'Perseverance of social theories: The role of explanation in the persistence of discredited information.', *Journal of Personality and Social Psychology* **39**(6), 1037.

Antweiler, W. & Frank, M. Z. (2004), 'Is all that talk just noise? the information content of internet stock message boards', *The Journal of finance* **59**(3), 1259–1294.

Araci, D. (2019), 'Finbert: Financial sentiment analysis with pre-trained language models', *arXiv preprint arXiv:1908.10063* .

Aral, S., Muchnik, L. & Sundararajan, A. (2009), 'Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks', *Proceedings of the National Academy of Sciences* **106**(51), 21544–21549.

Aral, S. & Nicolaides, C. (2017), 'Exercise contagion in a global social network', *Nature communications* **8**(1), 1–8.

Athey, S. & Imbens, G. W. (2017), 'The state of applied econometrics: Causality and policy evaluation', *Journal of Economic Perspectives* **31**(2), 3–32.

Banerjee, A., Chandrasekhar, A. G., Duflo, E. & Jackson, M. O. (2013), 'The diffusion of microfinance', *Science* **341**(6144).

Banerjee, A. V. (1993), 'The economics of rumours', *The Review of Economic Studies* **60**(2), 309–327.

Barberis, N., Greenwood, R., Jin, L. & Shleifer, A. (2018), 'Extrapolation and bubbles', *Journal of Financial Economics* **129**(2), 203–227.

Barlevy, G. & Veronesi, P. (2000), 'Information acquisition in financial markets', *The Review of Economic Studies* **67**(1), 79–90.

Bifulco, R., Fletcher, J. M. & Ross, S. L. (2011), 'The effect of classmate characteristics on post-secondary outcomes: Evidence from the add health', *American Economic Journal: Economic Policy* **3**(1), 25–53.

Black, F. (1986), 'Noise', *The Journal of Finance* **41**(3), 528–543.

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E. & Fowler, J. H. (2012), 'A 61-million-person experiment in social influence and political mobilization', *Nature* **489**(7415), 295–298.

Bordalo, P., Gennaioli, N., Kwon, S. Y. & Shleifer, A. (2021), 'Diagnostic bubbles', *Journal of Financial Economics* **141**(3), 1060–1077.

Bordalo, P., Gennaioli, N. & Shleifer, A. (2020), 'Memory, Attention, and Choice*', *The Quarterly Journal of Economics* **135**(3), 1399–1442.

Bouchaud, J.-P. & Potters, M. (2003), *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*, 2 edn, Cambridge University Press.

Bramoullé, Y., Djebbari, H. & Fortin, B. (2020), 'Peer effects in networks: A survey', *Annual Review of Economics* **12**, 603–629.

Breza, E. & Chandrasekhar, A. G. (2019), 'Social networks, reputation, and commitment: evidence from a savings monitors experiment', *Econometrica* **87**(1), 175–216.

Bulow, J. I., Geanakoplos, J. D. & Klemperer, P. D. (1985), 'Multimarket oligopoly: Strategic substitutes and complements', *Journal of Political Economy* **93**(3), 488–511.

Bursztyn, L., Ederer, F., Ferman, B. & Yuchtman, N. (2014), 'Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions.', *Econometrica* **82**(4), 1273.

Chen, H., De, P., Hu, Y. J. & Hwang, B.-H. (2014), 'Wisdom of crowds: The value of stock opinions transmitted through social media', *The Review of Financial Studies* **27**(5), 1367–1403.

Cookson, J. A. & Niessner, M. (2020), 'Why don't we agree? evidence from a social network of investors', *The Journal of Finance* **75**(1), 173–228.

Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. (2018), 'Bert: Pre-training of deep bidirectional transformers for language understanding', *arXiv preprint arXiv:1810.04805* .

Duflo, E., Dupas, P. & Kremer, M. (2011), 'Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya', *American Economic Review* **101**(5), 1739–74.

Epple, D. & Romano, R. E. (2011), Peer effects in education: A survey of the theory and evidence, *in* 'Handbook of social economics', Vol. 1, Elsevier, pp. 1053–1163.

Gabaix, X. & Koijen, R. S. (2020), Granular instrumental variables, Technical report, National Bureau of Economic Research.

Gabaix, X. & Koijen, R. S. (2021), In search of the origins of financial fluctuations: The inelastic markets hypothesis, Technical report, National Bureau of Economic Research.

Galaasen, S., Jamilov, R., Juelsrud, R. & Rey, H. (2020), Granular credit risk, Technical report, National Bureau of Economic Research.

Galeotti, A., Golub, B. & Goyal, S. (2020), 'Targeting interventions in networks', *Econometrica* **88**(6), 2445–2471.

Garber, P. M. (1989), 'Tulipmania', *Journal of Political Economy* **97**(3), 535–560.

Glaeser, E. L. & Nathanson, C. G. (2017), 'An extrapolative model of house price dynamics', *Journal of Financial Economics* **126**(1), 147–170.

Grossman, S. J. & Stiglitz, J. E. (1980), 'On the impossibility of informationally efficient markets', *The American economic review* **70**(3), 393–408.

Hellwig, C. & Veldkamp, L. (2009), 'Knowing what others know: Coordination motives in information acquisition', *The Review of Economic Studies* **76**(1), 223–251.

Hirshleifer, D. (2001), 'Investor psychology and asset pricing', *The Journal of Finance* **56**(4), 1533–1597.

Hirshleifer, D. (2020), 'Presidential address: Social transmission bias in economics and finance', *The Journal of Finance* **75**(4), 1779–1831.

Kumar, A. & Lee, C. M. (2006), 'Retail investor sentiment and return comovements', *The Journal of Finance* **61**(5), 2451–2486.

Lahno, A. M. & Serra-Garcia, M. (2015), 'Peer effects in risk taking: Envy or conformity?', *Journal of Risk and Uncertainty* **50**(1), 73–95.

MacKinnon, J. G. & White, H. (1985), 'Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties', *Journal of Econometrics* **29**(3), 305–325.

Patacchini, E. & Zenou, Y. (2016), 'Social networks and parental behavior in the intergenerational transmission of religion', *Quantitative Economics* **7**(3), 969–995.

Pool, V. K., Stoffman, N. & Yonker, S. E. (2015), 'The people in your neighborhood: Social interactions and mutual fund portfolios', *The Journal of Finance* **70**(6), 2679–2732.

Ross, L., Lepper, M. R. & Hubbard, M. (1975), 'Perseverance in self-perception and social perception: biased attributional processes in the debriefing paradigm.', *Journal of Personality and Social Psychology* **32**(5), 880.

Sabherwal, S., Sarkar, S. K. & Zhang, Y. (2011), 'Do internet stock message boards influence trading? Evidence from heavily discussed stocks with no fundamental news', *Journal of Business Finance & Accounting* **38**(9-10), 1209–1237.

Sacerdote, B. (2011), Peer effects in education: How might they work, how big are they and how much do we know thus far?, *in* 'Handbook of the Economics of Education', Vol. 3, Elsevier, pp. 249–277.

Securities & Commission, E. (1963), 'Report of special study of securities markets of the securities and exchange commission', *House Document* **95**(4).

Shiller, R. J. (1984), Stock prices and social dynamics, Technical Report 2, The Brookings Institution.

Shiller, R. J. (2005), *Irrational Exuberance: (Second Edition)*, Princeton University Press.

Shiller, R. J. (2014), 'Speculative asset prices', *American Economic Review* **104**(6), 1486–1517.

Shiller, R. J. (2017), 'Narrative economics', *American Economic Review* **107**(4), 967–1004.

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

Veldkamp, L. L. (2006), 'Media frenzies in markets for financial information', *American Economic Review* **96**(3), 577–601.

Voitalov, I., van der Hoorn, P., van der Hofstad, R. & Krioukov, D. (2019), 'Scale-free networks well done', *Physical Review Research* **1**(3), 033034.

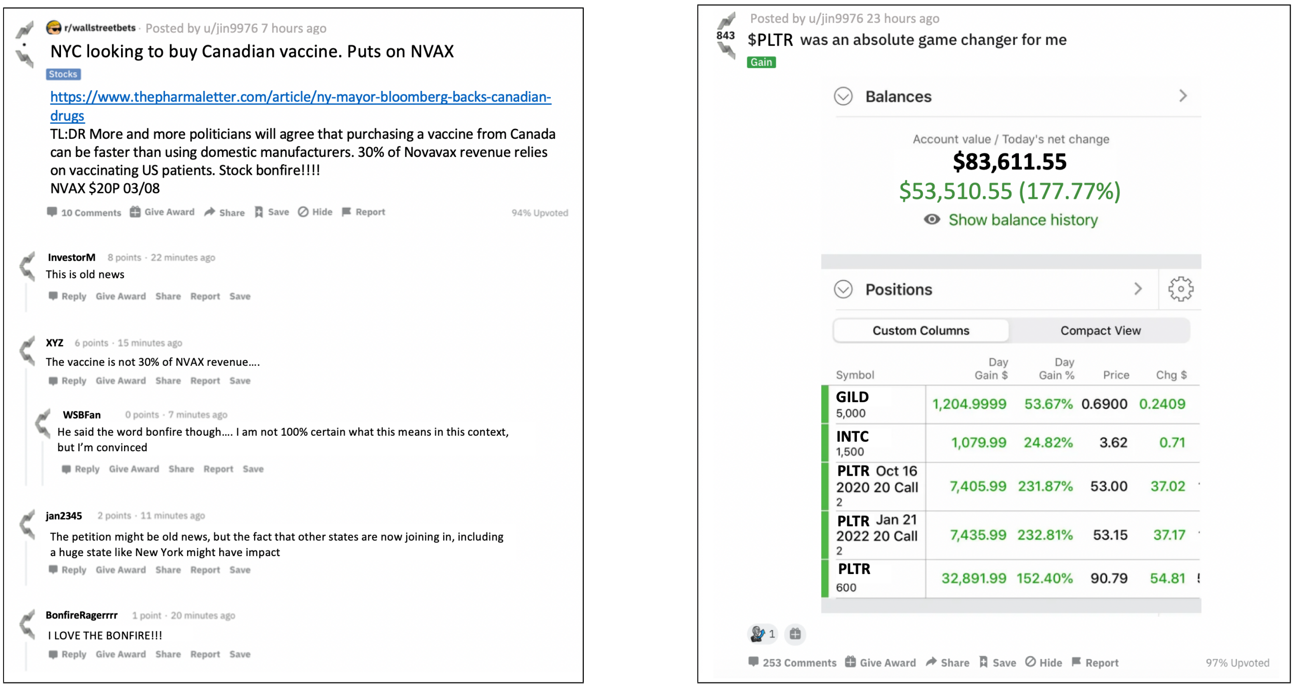
Weidmann, B. & Deming, D. J. (2021), 'Team players: How social skills improve team performance', *Econometrica* **89**(6), 2637–2657.

Zenou, Y. (2016), 'Key players', *The Oxford Handbook of the Economics of Networks* **11**.

Appendix

A Data appendix

A.1 Extended description of WSB



(a) A typical discussion on WSB

(b) A sample screenshot of user profits

Figure 6: **What does WSB look like?** These snapshots display typical discussions on WSB. The exact text, usernames, and conversation details have been modified to protect user identities.

Figure 6a displays a typical exchange on the WSB forum: individuals discuss stock-related news and their sentiments on whether this will affect stock prices in the future. In addition to market discussions, there is ample evidence of users pursuing the investment strategies encouraged in WSB conversations. Users post screenshots of their investment

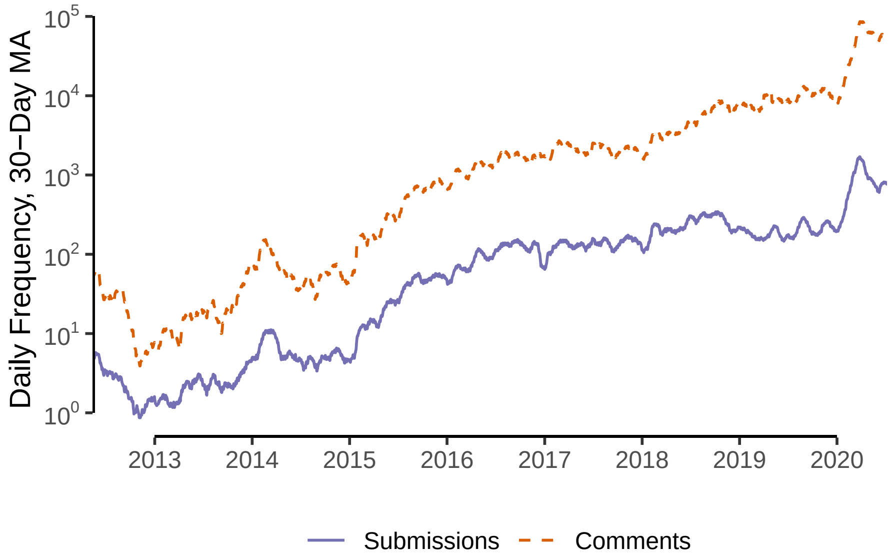


Figure 7: **Daily activity on WSB plotted on a logarithmic scale**; the daily submission and comment counts, averaged over 30 days, demonstrate a persistent exponential increase in activity on the WSB forum from 2015 to 2020, with a substantial jump in early 2020.

gains and losses, which subreddit moderators are encouraged to verify, as illustrated in Figure 6b.

Figure 7 displays the evolution of WSB over time. Two jumps are notable: a smaller, seemingly idiosyncratic rise in early 2018, and a sharp spike during the COVID-19 pandemic.

Reddit user content presentation Our empirical identification strategy in the *Frequent Posters* approach rests on the premises that users are exposed to random variation in peer sentiments. In this section, we discuss the details of how users are presented with content on Reddit.

Upon logging into Reddit, users are presented with a 'home feed'. Historically, the home feed has contained the 'top posts' from the subreddits to which a user has subscribed. More recently, Reddit has implemented an algorithm to try and match users to content based on a machine learning algorithm, via the home feed.[7] However, this change has only taken effect recently, and well after the point when our data were collected.[8] Subreddit top posts are not individually tailored to the specific user. Users have several sort options based on whether they prefer to see most recent or most highly rated content.[9] However, they do not have an option to personalize viewed content beyond this. We, therefore, conclude that individual users were exposed to WSB content based on the content on the forum that was most recent and popular at the time of their logging on, rather than based on personal preference. This, in turn, allows us to assert that users are exposed to random, temporal variation in peer sentiment.

[7] https://reddithelp.com/hc/en-us/articles/4402284777364-What-are-home-feed-recommendations-

[8] https://www.reddit.com/r/help/comments/rrkptm/home_feed_has_changed_drastically/

[9] https://www.reddit.com/r/help/comments/717686/order_of_posts/

A.2 Tickers mentioned on WSB

Table 8: Most frequent ticker mentions

Ticker	Name	Comments	Submissions	Sum
SPY	S&P 500 Index	291,279	9,408	300,687
AMD	Advanced Micro Devices, Inc.	124,685	5,721	130,406
TSLA	Tesla, Inc.	124,222	5,910	130,132
MU	Micron Technology, Inc.	86,611	3,941	90,552
AAPL	Apple Inc.	48,345	1,880	50,225
AMZN	Amazon.com, Inc.	44,426	1,534	45,960
MSFT	Microsoft Corporation	41,152	1,799	42,951
SNAP	Snap Inc.	40,766	2,043	42,809
NVDA	NVIDIA Corporation	38,012	1,556	39,568
SPCE	Virgin Galactic Holdings, Inc.	30,758	1,640	32,398
FB	Facebook, Inc.	26,143	1,446	27,589
DIS	The Walt Disney Company	25,611	1,088	26,699
BYND	Beyond Meat, Inc.	23,299	906	24,205
NFLX	Netflix, Inc.	20,800	936	21,736
JNUG	Direxion Daily Jr Gld Mnrs Bull 3X ETF	15,761	1,095	16,856
GE	General Electric Company	15,730	929	16,659
RAD	Rite Aid Corporation	14,781	839	15,620
SQ	Square, Inc.	14,003	824	14,827
ATVI	Activision Blizzard, Inc.	13,076	674	13,750
USO	United States Oil	12,949	667	13,616

Notes: this table lists the 20 most mentioned assets on WSB, observed by submissions which uniquely mention the related ticker. 'Comments' is the number of comments posted on these submissions, 'Submissions' counts submissions, and 'Total' is the sum of the two. The name of the asset corresponding to the identified ticker is retrieved from *Yahoo Finance*.

Conventionally, submissions or comments that mention a ticker will spell it using uppercase letters, or following a dollar sign. However, a challenge is that not all uppercase words are valid tickers.

We first match words in WSB submissions to assets by identifying any succession of two to five capital letters. Subsequently, we used a pre-determined list of tickers from CRSP to check whether a match is indeed present in the available financial data. Some abbreviations or capitalised words which are not valid tickers might still show up, such as 'USD' (*ProShares Ultra Semiconductors*), 'CEO' (*CNOOC Limited*), and 'ALL' (*The Allstate Corporation*). Single characters also appear, such as 'A' (*Agilent Technologies, Inc.*). We manually created a list of such tickers, and removed matches featured in WSB submissions, to build a preliminary list of candidate ticker mentions. We refined a second list of candidates by checking whether a collection of one to five letters, lower or uppercase, is preceded by a dollar sign. Any mentions of '$CEO' or '$a' count as the tickers 'CEO' and 'A', respectively. These extracts are, again, checked against the list of available tickers.

A small fraction of the 4,650 tickers we extract dominate the discourse on WSB. 90% of tickers are mentioned fewer than 31 times, and more than 60% are mentioned fewer than five times. The frequency distribution of tail of ticker mentions demonstrates this point,

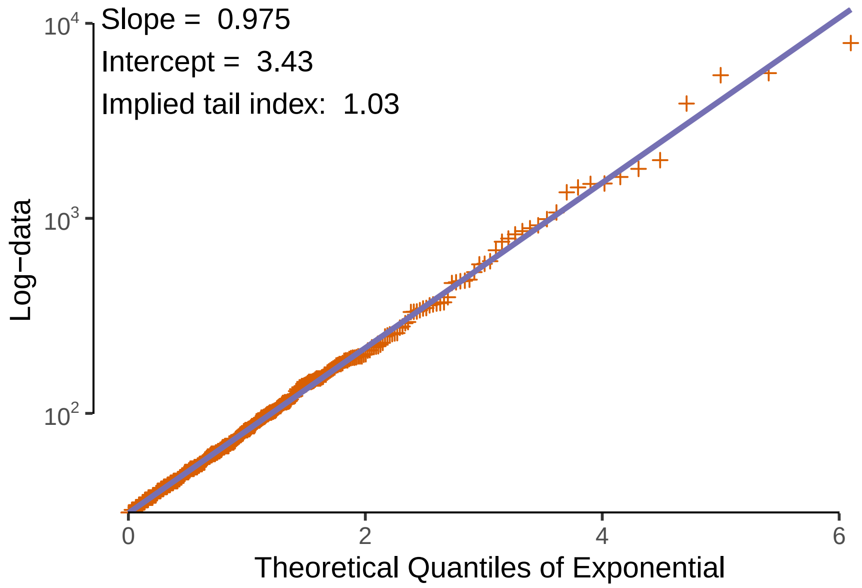


Figure 8: **QQ Plot of the tail in ticker mentions on WSB**; the number of submissions for each ticker (on a log-scale) is plotted against the theoretical quantiles of an exponential distribution. Quantiles are calculated as $q(i) = -\log(1 - i/(N+1)$, where N is the number of observations, and i the order of the statistic, from 1 to N. The linear fit suggests that the data follows a Pareto distribution, with the tail index equal to the inverse of the slope. The threshold for a ticker to be part of the 'tail' is 31 mentions; note the intercept, at $\exp(3.43) \approx 31$.

for which Figure 8 displays a QQ-plot. We arbitrarily selected tickers with the number of mentions in the top 10^{th} percentile. Even though threshold of mentions for this top decile is 30 submissions, the most popular, SPY, features in almost 8,000 submissions. The orange crosses in Figure 8 locate the empirical densities, on a log scale, which are plotted against the theoretical quantiles of an exponential distribution on the x-axis. Under the assumption that ticker mentions are heavy-tailed (similarly to vocabulary distributions), the logarithm of the mentions follows an exponential distribution, with the intercept at the threshold, and the slope equal to the inverse of the tail index. Indeed, the linear fit in Figure 8 is close to perfect, supporting the assumption that the popularity of assets in WSB is heavy-tailed, with an estimated tail exponent of approximately 1.03. In what follows, we used submissions for which we identified a single ticker, unless otherwise specified, forming a dataset of 103,205 submissions with unique ticker mentions by our cutoff date.

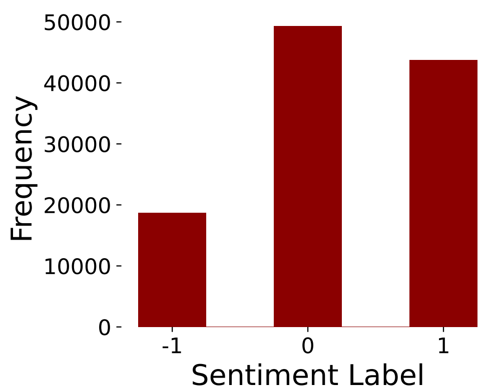
A.3 Sentiment modeling in WSB posts

True Label \ Predicted Label	-	**0**	+
-	64%	28%	7%
0	6%	77%	17%
+	6%	27%	67%

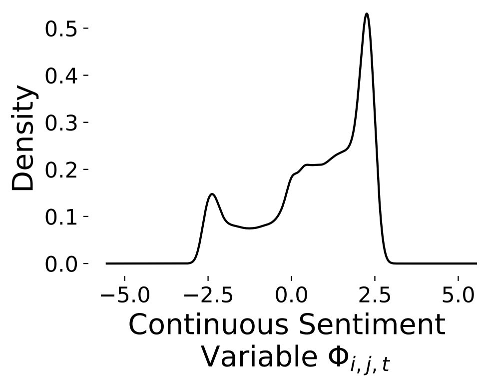
Table 9: **Fine-tuned FinBERT confusion matrix:** We use 10% of our hand-labeled data to test the performance of FinBERT on out-of-sample sentiment prediction. The results highlight the model's ability to predict sentiment with reasonably high accuracy.

Our goal, with regards to the text data in WSB, is to gauge whether discussions on certain assets express an expectation for their future price to rise, the 'bullish' case, to fall, the 'bearish' case, or to remain unpredictable, the 'neutral' case. Among other alternatives, we pursued a supervised-learning approach to identify the sentiment expressed about an asset within a WSB submission. This required a training dataset, for which we manually labelled 4,932 random submissions with unique ticker mentions as either 'bullish', 'bearish' or 'neutral' with respect to the authors' expressed expectations for the future price. We used the FinBERT algorithm for labeling (Araci 2019) - a financially oriented modification of Google's Bidirectional Encoder Representations from Transformers (BERT) algorithm (Devlin et al. 2018). Work not shown here implements an alternative regression-based approach as a robustness check, but FinBERT performs better out-of-sample.

We trained FinBERT on 75% of the labelled data, and used the remaining 25% for validation and the test set. Table 9 plots the out-of-sample confusion matrix. For the out-of-sample test, we train FinBERT on 75% of the available data and use 15% for validation; we then compute what the algorithm predicts for the remaining 10% of data. We achieve 70% accuracy on the test set. This is better than a LASSO regression's accuracy, which was implemented separately and is not cover here.



(a) Distribution of Sentiment Labels - Output from FinBERT Classifier



(b) Continuous Sentiment Variable Distribution

Figure 9: **Distribution of Expressed Sentiments on WSB**; We present the density plot of the labeled posts on WSB, and our key continuous variable of log-odds of positive over negative sentiment $\Phi_{i,j,t}$.

Data Description Our final samples contains 111,765 submissions that have a mention of a single, identifiable asset. Figure 9a shows that the sample is slightly unbalanced - more posts are labeled as neutral than the other two categories, and more posts appear bullish than bearish. The distribution of our continuous variable is shown in Figure 9b.

A.4 Market variables

We include a set of market return and volatility control variables. The data source for these variables are the daily stock files issued by the Center for Research in Security Prices (CRSP), accessed through Wharton Research Data Services.

Market variables in Sections 2.2&4.1 The following market variables serve as controls.

$r_{j,t}$: the log return for asset j on trading day t. From CRSP, we calculate it using their 'RET' variable: $r_{j,t} = \log(RET_{j,t} - 1)$, which automatically corrects the percentage change in closing prices for share splits and dividend distributions.

$\bar{r}_{j,t}$: the average log returns for asset j in the five days prior to t (the log return on day t is not included). A minimum of three daily log-return observations is required, otherwise the observation is set as missing.

$\sigma^2_{j,t}$: the variance of log returns for asset j in the five days prior to t (the log return on day t is not included). A minimum of three daily log-return observations is required, otherwise the observation is set as missing.

Matching submission timings to trade timings If a post occurs before 16:00:00 EST on day t, we match it with the log-return on the same day t. If a post occurs after 16:00:00 EST on a given day, we match it with market data for the next trading day, $t+1$. This is done to capture the fact that many news announcements occur after hours and someone posting after the market close may be exposed to these after-hour moves. Instance in which submissions are made on weekends, or holidays, are matched to the next possible trading day. For example, a submission made at 5pm on Friday is paired to the observed log return for the following Monday.

B Extra results for peer effects

B.1 Target independent variable

We build a discrete-choice empirical strategy to suit our model. Under the assumption that $u_{i,t}$ is drawn from a standard type-I Extreme Value Distribution, we model the log-odds of expressed investor sentiments ϕ_i by a standard multivariate logistic function,

$$\log\left(\frac{P(\phi_{i,t} = +1)}{P(\phi_{i,t} = 0)}\right) = g(b_{i,t}) + f(\bar{\phi}_{-i,(t-1,t)}) - \theta\sigma^2_{i,t} + u^+_{i,t}, \tag{33}$$

$$\log\left(\frac{P(\phi_{i,t} = -1)}{P(\phi_{i,t} = 0)}\right) = -g(b_{i,t}) - f(\bar{\phi}_{-i,(t-1,t)}) - \theta\sigma^2_{i,t} + u^-_{i,t}, \tag{34}$$

where t denotes time, and $(t-1,t)$ an interval preceding t. The goal of this paper, in light of Prediction 1, is to test empirically whether $f(\cdot)$ is increasing. To that end, we aggregate bullish and bearish sentiments into one continuous variable, $\Phi_{i,t}$:

$$\Phi_{i,t} = \frac{1}{2}\log\left(\frac{P(\phi_{i,t}=+1)}{P(\phi_{i,t}=-1)}\right) = g(b_{i,t}) + f(\bar{\phi}_{-i,(t-1,t)}) + \frac{u^+_{i,t} - u^-_{i,t}}{2}. \tag{35}$$

In the main body, the error term is expressed as $\epsilon_{i,t}$. Under the assumption that $u^+_{i,t}$ and $u^-_{i,t}$ are independent and identically distributed, $u^+_{i,t} - u^-_{i,t}$ will follow a logistic distribution with finite variance.

B.2 Full regression estimates

Tables 10 and 12 present our full regression estimates. Table 11 presents our First Stage estimates for our *Commenter Network* approach, which has multiple IVs.

Table 10: Peer influence: *Frequent Posters* – full regression estimates

		Dependent Variable: $\Phi_{i,j,t}$	
	Reduced Form (1)	Full Second Stage (2)	Random Peers (3)
Independent Variables $\Phi_{i,j,t-1}$	0.15 (0.01) ***	0.13 (0.01) ***	0.15 (0.01) ***
$\bar{\Phi}_{-i,j,(t-1,t)}$	0.10 (0.02) ***	0.19 (0.05) ***	0.01 (0.02)
$r_{j,t}$	0.85 (0.13) ***	0.95 (0.17) ***	0.89 (0.14) ***
$\bar{r}_{j,t}$	0.80 (0.42) *	0.87 (0.50) *	0.82 (0.44) *
$\sigma^2_{j,t}$	-0.39 (0.61)	0.24 (1.21)	-0.38 (0.61)
Ticker Fixed Effects	Yes	Yes	Yes
No. Observations:	14,396	11,122	14,391
R^2:	0.12	0.08	0.11
R^2_{adj}:	0.08	0.06	0.08

Notes: The dependent variable is individual investor sentiment about an asset, scaled continuously between $(-\infty,\infty)$, is estimated by the individual's previously expressed sentiment about the same asset ($\Phi_{i,j,t-1}$) and a set of market control variables ($r_{j,t}, \bar{r}_{j,t}, \sigma^2_{j,t}$), using OLS. The sentiment of peers ($\bar{\Phi}_{-i,j,(t-1,t)}$) is estimated in several ways. In Column (1), we use observed, average sentiment of peers between an author's two posts. In Column (2), we estimate the sentiment of peers using an IV. In Column (3), we select a random cohort to estimate peer sentiment. Robust standard errors, clustered at the ticker level, are presented in parentheses. Observations with incomplete market data are dropped.
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

B.3 Evidence of identification strategy

A potential concern with our approach is whether the sentiments expressed by individuals who post multiple times or are part of the commenters network follow the same distribution as all submissions on the forum. Figure 10a presents the distribution of sentiments for the second or later post of an author about a ticker and Figure 10b presents the distribution of sentiments for those who comment on other's posts. Figure 10 provides evidence that

Table 11: First Stage estimates for *Commenter Network* approach

	$\phi^{-1}_{i,j,t-1}$	$\phi^{0}_{i,j,t-1}$	$\phi^{+1}_{i,j,t-1}$	$\bar{\Phi}_{-i,j,t-1}$
Dependent variable:				
Sentiment of Peers	-0.30 (0.04) ***	0.12 (0.03) ***	0.25 (0.03) ***	0.14 (0.01) ***

Notes: The dependent variable is individual investor sentiment about an asset expressed in a single submission, scaled continuously between $(-\infty, \infty)$, modeled using IVs. We estimate it using the individual's previously expressed sentiment about the same asset ($\phi_{i,j,t-1}$) as a categorical variable, with the author not having posted previously ($\phi^{NA}_{i,j,t-1}$) as the baseline, as well as the average sentiment of posts that the author commented on previously ($\bar{\Phi}_{-i,j,t-1}$). We user the timing of IVs to control for common shocks, as discussed in the main text. Our regression has 24,013 observations and an F-statistic of 118.
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 12: Peer influence: *Commenter Network* – full regression estimates

		Dependent Variable – $\Phi_{i,j,t}$	
	Reduced Form	Full Second Stage	Random Network
	(1)	(2)	(3)
Independent Variables			
$\phi^{-1}_{i,j,t-1}$	-0.34 (0.04) ***	-0.32 (0.03) ***	-0.35 (0.04) ***
$\phi^{0}_{i,j,t-1}$	0.07 (0.03) **	0.05 (0.03)	0.07 (0.04) **
$\phi^{+1}_{i,j,t-1}$	0.24 (0.04) ***	0.21 (0.05) ***	0.25 (0.04) ***
$\bar{\Phi}_{-i,j,t-1}$	0.05 (0.01) ***	0.19 (0.08) **	0.01 (0.01)
$r_{j,t}$	0.84 (0.12) ***	0.98 (0.24) ***	0.85 (0.12) ***
$\bar{r}_{j,t}$	0.90 (0.42) **	1.31 (0.73) *	0.95 (0.43) **
$\sigma^2_{j,t}$	-0.01 (0.51)	0.57 (0.95)	0.00 (0.53)
Ticker Fixed Effects	Yes	Yes	Yes
No. Observations:	24,963	16,521	25,284
R^2:	0.09	0.07	0.09
R^2_{adj}:	0.06	0.06	0.06

Notes: The dependent variable is individual investor sentiment about an asset expressed in a single submission, scaled continuously between $(-\infty, \infty)$. We estimate it using the individual's previously expressed sentiment about the same asset ($\phi_{i,j,t-1}$) as a categorical variable, with the author not having posted previously ($\phi^{NA}_{i,j,t-1}$) as the baseline. We control for a set of market control variables ($r_{j,t}, \bar{r}_{j,t}, \sigma^2_{j,t}$). The sentiment of posts that the author commented on previously ($\bar{\Phi}_{-i,j,t-1}$) is estimated several ways. In column (1), we present the estimate using the sentiment of posts the author previously commented on. In column (2), we use an IV to predict the sentiment of posts the author comments on. In column (3), we randomly rewire the network, connecting the author to a random set of posts about the same ticker. Robust standard errors, clustered at the ticker level, are presented in parentheses. Observations with incomplete market data are dropped.
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

the sentiment distributions are similar to that of other posters on WSB, which supports the hypothesis that our analysis offers insight into how all individuals on WSB form opinions.

A second concern is whether we effectively control for unobserved ticker characteristics. Similarly to Patacchini & Zenou (2016), we run 'placebo tests', where we replace the composition of an author's peers with a random cohort of people who post on WSB about the same ticker. The random cohort is chosen as follows. We observe how many peers an individual author has. We then select a random sample of the same number of individuals, without replacement, who do not post between an author's two post but post about the ticker at a

(a) Frequent Posters Sentiment PDF

(b) Commenters Sentiment PDF

Figure 10: **Density Plot of Sentiments Expressed on WSB**; We present the density plot of the sentiments expressed by users on WSB who post multiple times, labeled as *Multiple Posters*, those who comment on others' posts, labeled as *Network*, and that of all submissions, labeled as *All Posts*.

different time for the *Frequent Posters* approach (if fewer individuals post before, we select all of those individuals), or through a random network rewiring (we select posts randomly about the same ticker before the current post). The results are presented in Tables 10 and 12, column (3). We observe that all the coefficients remain close to their original values, except for the peer effect, which becomes insignificant. This lends credibility to our peer identification strategy and shows that unobserved factors that influence within ticker variation are not confounding our estimates.

We cannot directly calculate the J-statistic for our *Commenter Network* approach, since we estimate our IV using observations on several neighbours. We, therefore, take an average of the neighbours past sentiments (transforming the categorical variable into a continuous one) and the average across their neighbour's neighbours sentiments. We use this to compute a J-Statistic with two degrees of freedom.

B.4 Contagion dynamics and the origin of bull runs

In the WSB context, we would expect awareness about specific assets to spread from one user to another, in line with the observations of Shiller (2005), Banerjee (1993) and Banerjee et al. (2013). The emphasis of this section is not on identifying a causal relationship, but rather understanding the dynamics which govern asset interest among investors. These insights, combined with a mechanism for investors' joint sentiment adoption, allow us to paint a more complete picture of retail investor decision-making and the resultant stock market dynamics.

We model the log-odds of an author posting about stock j over a baseline using the following linear model:

$$l(a_{j,t}) = \log\left(\frac{a_{j,t}}{s_t}\right) = c a_{j,t-1}(1-a_{j,t-1}) + d a_{j,t-1} + \beta_1 \bar{r}_{j,t-1} + \beta_2 \sigma^2_{j,t-1} + X_j \beta_4 + \zeta_{j,t}, \tag{36}$$

where t denotes time (in weeks), the baseline s_t is the probability of posting about a stock that is not widely discussed within the forum (a stock that is mentioned in fewer than 31 submissions within our sample), $a_{j,t-1}$ is the share of all active investors who post about ticker j at times $t-1$ ($a_{j,t} \in [0,1]$ for all j and t), $\bar{r}_{j,t-1}$ is the average log-return in $t-1$, and $\sigma^2_{j,t-1}$ is the variance of the same log-returns (these variables are mostly consistent with Section 4.1, and discussed further in our Online Appendix). X_j is a vector of stock dummies.

Our framework resembles that of Section 4.1 and is inspired by Banerjee et al. (2013) – individuals become interested in an asset due to their peers and thanks to a public signal of the asset's performance. Parameter c captures the rate of independent mixing between investors aware of stock j, $a_{j,t-1}$, with unaware investors, $1-a_{j,t-1}$. Parameter d captures the rate at which aware investors lose interest. The latter terms control for the asset's perceived profitability and riskiness. Parameter β_1 is a 'quality of signal' term capturing how well the asset has performed in the past, and β_2 a 'noise of signal' term, measuring the asset's recent volatility. We propose that coefficients c and β_1 are positive – implying that these dynamics contribute to increased interest in a stock – while d and β_2 are negative.

The choice to aggregate over weeks is done to address the sparsity of submissions, especially pre-2017. In addition, we categorise stocks mentioned fewer than 31 times since January 2012 into an 'other stocks' group, which forms our benchmark s_t.

We also consider a different formulation where we test for the direct impact of historical peer sentiments and the interactions between historical sentiments and returns / volatility: $\bar{\phi}_{j,t-2}\bar{r}_{j,t-1}$, $\bar{\phi}_{j,t-2}$ and $\bar{\phi}_{j,t-2}\sigma^2_{j,t-1}$. This formulation allows us to evaluate whether WSB users are more likely to discuss a stock if the predictions of their peers have been correct, and accurate, in the past.

B.4.1 Results

The OLS estimates of our model in Eq. 36, presented in Table 13, demonstrate that WSB users follow each other in their choice of stocks. There is strong evidence that the homogeneous mixing property partially explains the uptake of new assets: using estimates in column (1), an increase in the share of authors discussing stock j from 0.1 to 0.2 increases the ratio of authors discussing j over 'other stocks' in the following week by approximately threefold. This is contrasted by an increase from 0.2 to 0.3, which prompts a decline in the ratio of authors discussing j over 'other stocks' by 50% – the difference is driven by the large negative coefficient on $a_{j,t-1}$. This is strongly reminiscent of epidemic contagion models, adapted to the spread of narratives (Banerjee 1993, Shiller 2017).

When we consider the impacts of stock-specific variables in isolation, presented in columns (1), (3) in Table 13, volatility and returns appear to be leading factors for authors deciding what asset to discuss. Average, historical returns are statistically significant at the 1% level in columns (1) and (3), indicating that discussion sizes are stimulated by large, notably positive, returns. Examining the coefficient in column (3), a stock that experienced a 5% greater return in one week is the subject of about 7% more submissions than usual. Volatility appears to play a greater role in our formulation without ticker-specific effects, with its significance declining from column (1) to (3) – a factor perhaps explained by the choice of hype investors to overlook recent volatility in certain assets, but not others. Our alternative formulation presented in columns (2) and (4), estimating the effect of the correctness

Table 13: Stocks discussed on WSB

	Dependent variable: $l(a_{j,t})$			
	(1)	(2)	(3)	(4)
$a_{j,t-1}(1-a_{j,t-1})$	83.49*** (8.20)	100.20*** (9.15)	46.30*** (5.33)	57.90*** (5.13)
$a_{j,t-1}$	−48.01*** (7.04)	−62.37*** (7.83)	−24.06*** (3.94)	−33.73*** (3.95)
$\bar{r}_{j,t-1}$	1.24*** (0.39)		1.36*** (0.42)	
$\sigma^2_{j,t-1}$	−2.15*** (0.60)		−0.96* (0.54)	
$\bar{\phi}_{j,t-2}\bar{r}_{j,t-1}$		0.56 (1.09)		1.59 (1.10)
$\bar{\phi}^2_{j,t-2}\sigma^2_{j,t-1}$		−5.14** (2.19)		−1.71 (1.53)
Constant	−3.89*** (0.01)	−3.88*** (0.02)		
Ticker FE	No	No	Yes	Yes
Number of obs.	13,184	6,429	13,184	6,429
Adjusted R^2	0.28	0.36	0.10	0.14
F-statistic	1,294	920	429	318

Notes: this table presents OLS estimates for the log-odds of users discussing stock j in week t, over a collection of stocks that are mentioned fewer than 31 times. Explanatory variables include: the lag in the share of authors discussing j, $a_{j,t-1}$, the interaction with the share of authors not discussing j, $a_{j,t-1}(1-a_{j,t-1})$, as well as the lag in stock j's weekly average log-return, $\bar{r}_{j,t-1}$, and variance, $\sigma^2_{j,t-1}$. In columns (2) and (4), the average log-return is multiplied by the two period lag in the average sentiment expressed among WSB submissions on stock j, $\bar{\phi}_{j,t-2}$, and the variance in log-returns by the same sentiment's square, $\bar{\phi}^2_{j,t-2}$. Columns (3) and (4) include stock-specific fixed effects. Accompanying standard errors, displayed in brackets, are clustered at the stock level, and calculated in the manner of MacKinnon & White (1985).
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

and consistency of past WSB predictions in an asset, appears to have limited significance in explaining asset interest.

C Extra results for market impact

C.1 Portfolio

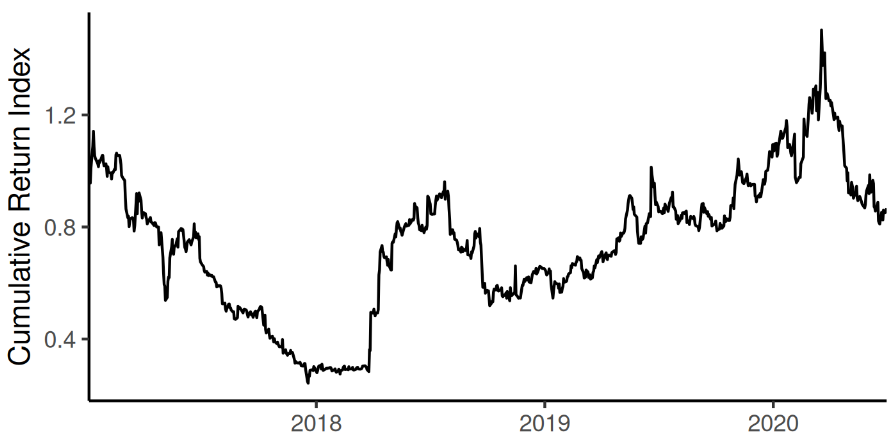


Figure 11: **Cumulative returns of a portfolio built on WSB sentiment.**

The performance of a portfolio that buys stocks according to WSB sentiments is questionable. Among alternative transformations, we assign weights for stocks with net positive sentiment, $w+_t$, and negate sentiment w_t^-. These proxy for long and short positions, with the size of the position by stock measure by the ratio of the stocks' lagged sentiment over the sum of sentiments:

$$w_{i,t}^+ = \frac{\Phi_{i,t-1}^+}{\sum_i^N \Phi_{i,t-1}^+}, \quad w_{i,t}^- = \frac{\Phi_{i,t-1}^-}{\sum_i^N \Phi_{i,t-1}^-}. \tag{37}$$

The cumulative return of the combined long/short portfolio is displayed in Figure 11. The indexed return from the start of 2017, when WSB data turned frequent enough to trade on daily, to July 2020 displays losses amounting to %13.3. These are punctuated by various periods of consistent gains or losses, plus certain days of large outsized returns. The information content on WSB therefore does not appear inherently valuable, although this is admittedly not the stated intent; the volatile outcomes speak to the users' aspirations for large, one-off gambles.

C.2 GIV

Extracting idiosyncratic social shocks In Table 14, we present the coefficients from Eq. 32. We observe that, consistently with our previous findings and the proposed structure in Figure 2, returns and past user sentiments all impact the expressed sentiment of user i about asset j at time t.

Table 14: Idiosyncratic social shocks estimation

	Dependent variable: $\Phi_{i,j,t}$
$r_{j,t}$	0.327*** (0.099)
r_{j,t_w}	0.214* (0.085)
$\bar{\Phi}_{j,t_w}$	0.250*** (0.000)
Ticker FE	Yes
Number of obs.	38,060
R^2	0.076
F-statistic	28.77

Notes: The dependent variable is the log-odds of a given submission by author i at time t on stock j to express bullish over bearish sentiment. Explanatory variables include: the log return on day t, $r_{j,t}$, the log-return in week t_w, r_{j,t_1}, and the average past sentiment of peers, $\bar{\Phi}_{i,j,t_w-1}$. Accompanying standard errors, displayed in brackets, are clustered at the stock level.
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

GIV - First Stage In Table 15, we present our First Stage regression, where we regress the popularity-weighted idiosyncratic sentiment, as our dependent variable, on the difference between the popularity-weighted and regular average idiosyncratic sentiment. We observe that our GIV is highly predictive of the popularity-weighted sentiment, $\bar{e}_{j,t_w}$, however, explains only part of the variation. In this way, we are able to extract the element of the popularity-weighted sentiment measure driven by social preferences, versus those driven by other factors.

Table 15: GIV - First Stage

	Dependent variable: $\bar{e}_{j,t_w}$
GIV	0.946*** (0.099)
Number of obs.	916
R^2	0.575
F-statistic	1239

Notes: The dependent variable is the popularity-weighted average idiosyncratic sentiment expressed about asset j in week t_w, $\bar{e}_{j,t_w}$. The explanatory variable is the difference between the popularity-weighted and raw average of idiosyncratic sentiments expressed about asset j in week t_w.
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

We illustrate the effect of the GIV through the following two scenarios. In *Scenario One*, a very popular and an unpopular post both express a positive, idiosyncratic sentiment. In *Scenario Two*, a very popular expresses the positive sentiment, while the unpopular posts expresses no idiosyncratic sentiment. Without employing the GIV, our original popularity-weighted average idiosyncratic sentiment measure $\bar{e}_{j,t_w}$ would be very similar in both scenarios. However, by using the GIV, we would *predict* zero idiosyncratic sentiment in *Scenario One*, but a large positive idiosyncratic sentiment in *Scenario Two*. In this way, our GIV allows us to distinguish the signal coming from popularity.

GIV requirements and threats to identification In addition to capturing the impact of social forces, our specification allows us to disentangle asset properties which affect both sentiment and returns simultaneously. Specifically, stock j's returns in week t_w+1 may be driven in-part by social forces, but also by asset properties which affect both sentiment and returns, which we are unable to control for while estimating idiosyncratic sentiments in Eq. 30. A correlation of these shocks with $\bar{e}_{i,t_w}$ may result for a biased estimator for β_1^B. More formally, outcome variable (after imposing controls from Eq. 31) y^r_{j,t_w+1} may be of the form:

$$y^r_{j,t_w+1} = \beta_1^B \bar{e}_{j,t_w} + \eta_{j,t_w}, \tag{38}$$

where η_{j,t_w} is the 'common shock' to asset j in period t_w.

We assume that the idiosyncratic social shock from a post can be expressed as having a stock level component, common to all posts about the stock within that time period, and a post-level component:

$$e_{i,j,t_w} = \beta^{cs}_{j,t_w} \eta_{j,t_w} + u_{i,j,t}, \tag{39}$$

where β^{cs}_{j,t_w} is the sensitivity of posts within week t_w to the common shock to stock j.

The identification strategy rests on the assumption that the popularity of the idiosyncratic content of posts $s^c_{i,j,t_w} e_{i,j,t_w}$ is not correlated with stock-week shocks η_{j,t_w}. More formally, we require $E[u_{i,j,t}\eta_{j,t_w}] = 0$: the idea is that there are social shocks which make certain content on WSB popular over other content, but that is orthogonal to shocks affecting asset j in week t_w. This is not a problem in this setting for several reasons. Firstly, the popularity of a post could potentially be linked to a stock through it's informativeness about the asset's price. However, in the creation of our social idiosyncratic shocks, we extract shocks while controlling for returns on the day and the week of the post. Our social shock time series is, therefore, orthogonal to asset returns at time t and in week t_w and is, therefore, orthogonal to new information available to investors at the time. Furthermore, we find post popularity to be uncorrelated with asset returns on day t and week t_w on which the post is made. Second, in WSB we observe data about relatively unsophisticated retail investors where the sentiments of posts at time t about an asset are systematically linked to *negative* future returns (this holds both when we take a raw average and popularity-weighted average average sentiment). Therefore, the investors we observe do not have access to information on stock-level shocks. Finally, both our shock and popularity time series at constructed at time t_w, while the dependent variable is observed at time $t_w + 1$, avoiding contemporaneity issues.