Reddit's Self-Organised Bull Runs

Semenova, Valentina and Winkler, Julian

Institute of New Economic Thinking at the Oxford Martin School, Department of Mathematics, University of Oxford, Department of Economics, University of Oxford

20 January 2020

Reddit's Self-Organised Bull Runs

Valentina Semenova[1,2] and Julian Winkler[1,3]

[1] *Institute of New Economic Thinking at the Oxford Martin School*

[2] *Department of Mathematics, University of Oxford*

[3] *Department of Economics, University of Oxford*

January 20, 2021

Abstract

Tracking endogenous fluctuations in stock prices emerged as a key challenge for empirical work in behavioural and evolutionary finance. This paper uses new data from an online discussion forum, Reddit, to quantify social contagion, or 'hype,' in specific stock market movements, using state of the art opinion dynamics modelling and sentiment analysis. The influence between users on the WallStreetBets (WSB) subreddit is measured by tracing the probability of a user starting a fresh discussion on an asset given their previous involvement in a discussion on the same asset, measured by their comment history. This paper finds that users who comment on one discussion involving a particular asset are approximately four times more likely to start a new discussion about this asset in the future, with the probability increasing with each additional discussion the user engages in. This is a strong indication that investment strategies are reproduced through social interaction. This is further validated by findings that sentiments expressed in the linked submissions are strongly correlated in a set of spatial regression models. In particular, bearish sentiments seem to spread more than their bullish counterparts.

JEL codes: G14, G41.

Acknowledgements

We thank Rick Van der Ploeg, J. Doyne Farmer and Steve Bond for supervising this research, and Xiaowen Dong for his excellent comments and advice. We are grateful to John Pougué Biyong for pointing us to BERT, and Matthias Winkler for bringing our attention to hedge funds' excitement in prowling Reddit. 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.

Contents

1 Introduction

Recent efforts in behavioural and complexity economics emphasise the role of social interaction and narratives in driving asset price fluctuations. However, difficulties in gauging investor's social behaviour on a granular scale poses a significant empirical challenge to measuring these factors.

Social media presents an answer. The emergence of the forum named 'WallStreetBets' (WSB) underscores a rise of online forums where anonymous users are encouraged to discuss, share and buy into high-risk positions. Enabled by the expanding availability of online trading platforms, these social feeds offer an opportunity to track evolving opinions about assets, from their inception to widespread consensus formation, as non-institutional traders coordinate strategies.

This paper leverages the rich text data found within WSB using Natural Language Processing (NLP) techniques and Opinion Dynamics (OD) models. NLP has recently added traction to a literature that seeks to gauge investor sentiment, but the inherent interest in such information meant that early efforts date as far back as Cowles (1933). Bollen et al. (2011) and Manela & Moreira (2017), among others, apply dictionary and regression methods to predict stock prices and volatility from text data. These are typically sourced from mainstream news or social media feeds, but alternatives are plentiful. Gentzkow et al. (2019) write a thorough review on the state-of-the-art, yet the field is evolving rapidly.

The first part of this paper offers qualitative insights into the behaviour of non-institutional investors. The observed discussions display three key attributes. First, trades encouraged on the platform are typically high-risk gambles in options markets. The striking nature of these trades thus stands out as an aberration in traditional portfolio theory. Second, users pursue a mix of narratives, sometimes sarcastic in tone, in addition to offering or seeking trading advice. In doing so, a handful of narratives, estimated by a topic model, consistently follow real world trends, such as the American election, the legalisation of marijuana, and the COVID-19 pandemic. Third, a few assets dominate the conversation. In fact, the frequency distribution of submissions and associated comments is heavy tailed, with a tail index around 0.9. More extreme than typical word frequency distributions, a few assets do not just dominate overall discourse, but grow to define it.

Subsequently, this paper explores the intuition behind opinion formation through social interaction. Opinions are measured using an advanced neural net, with the outer layer trained on a dataset of 2,774 bearish, neutral or bullish WSB texts. The model's accuracy is satisfactory, and outperforms an alternative penalised regression model out of sample. Similar to previous studies, these sentiments are economically and statistically significant for asset price movements; bullish submissions correlate with positive excess returns, whereas the converse applies for bearish submissions. Market-adjusted log trading volumes and high-low price gaps are also found to be significantly higher on days where submissions are made, which is unsurprising given the forum's exigence for risky strategies. The observed opinions are thus certainly informed and reflective of the real world.

The goal is to reconcile the extreme nature of the behaviours in WSB with economic theory. The two aberrant characteristics are that users i) are encouraged to gamble, with examples of authors displaying their oft collapsing, but occasionally ballooning balances, and ii) galvanise their peers to adopt similar positions. One economic theory to rationalise this behaviour is the 'narrow framing' coined in Barberis & Huang (2009). Users are exposed to specific positions from discussions they participate in, and are more likely to reproduce these strategies in isolation, instead of evaluating them against other possibilities. The observed pattern of contagion for demand of assets then strongly follows what would be expected from a typical preferential attachment model, by which the probability of adoption for a strategy increases with its centrality in the WSB discourse.

The working hypothesis is that user's expressed interest in assets, our proxy for asset demand, depends on who they engage with. The aim is to measure the propensity to post a submission on an asset conditional on past participation in a discussion mentioning this asset. This is sensible to the extent that WSB's interest in an asset is effectively dead if no submissions appear about it. Accurately evaluating such adoption is not straightforward; endogenous link formation may strongly influence the associated spillovers, as highlighted in Aral et al. (2009). In the context of this paper, this limitation is significant since users self-select in posting on WSB. Unobserved characteristics thus bias the estimates of a naive network model. This issue is addressed with an OD model that matches users on observable characteristics, a method adopted from Leng et al. (2018). The observable characteristics are drawn from the vast data on user activity on the whole Reddit website. The results demonstrate significant contagion in asset demand after accounting for endogeneity. Users who comment in discussions about an asset are between four and nine times more likely to subsequently start a new conversation about the asset themselves, compared to their matched counterparts in the control group.

The next question is whether predictions of asset performance, or sentiments, also transmit between users. Given the robustness of the discussion networks to unobserved heterogeneity, the spread of sentiments on an asset is well measured by the associated spillover from the network of discussions. This network links submissions by their authors' comments on older submissions mentioning the same asset. The results indicate that sentiments in those submissions correlate strongly and significantly with the sentiments of neighbours, especially for bearish post; the submission from a user who commented on a bearish post in the past is 65.8% more likely to be bearish than indifferent. The conclusion drawn is that social interaction plays an important role in determining investor sentiment, net of actual fundamentals, but this asymmetry is of particular interest.

The structure of the paper is as follows. Section 2 contextualises this research by reviewing existing papers on behavioural finance, NLP in economics and OD. Section 3 comprehensively describes the data, plus the steps taken in extracting the relevant variables. It reports on the results of a topic model to describe the key narratives underpinning the observed discussions. It also offers some summary statistics on the featured assets, as well as their associated sentiments. Section 4 estimates the contagion of both interest in assets and their associated sentiments. Section 5 concludes, and offers goals for further research.

2 Literature Review

This section offers context to this paper by reviewing the pertinent studies that use text data in finance and opinion dynamic tools to model social interaction. In doing so, it presents some existing theory from behavioural finance to understand the economic motivations behind the observed behaviours in WSB. Subsequently, section 2.2 contextualises the approach of this paper in the existing body of work in finance that leverages text data. This leads into a short review on current approaches in OD, largely stemming from budding areas in mathematics and computer science, in section 2.3.

2.1 Gambles and Herds

The best context to place the behaviour observed in WSB is within the results of Barberis et al. (2006). The authors explore the historically difficult issue of reconciling functional approaches for economic agent's risk preferences with the empirical observation that they choose large gambles, rather than small yet favourable ones. They address this issue with some success by introducing a mechanism for 'narrow framing,' by which gambles are evaluated independent of outstanding wealth and their associated risks.

Intuitively, as suggested in the paper, investment decisions are governed by the prospect of regret, should a gamble fail, or fear of missing out, should a gamble indeed succeed. This aligns well with qualitative evidence that users in WSB often enter high risk positions advertised by others, specifically for fear of missing out on profits. Interesting dynamics among investors are already touched upon in Farmer (2002), positing that profits can be made by trend followers who convince peers to adopt their strategy. Evidence of such behaviour is documented in Musciotto et al. (2018), who track cluster of investors using their trading profiles. They shed light on the heterogeneity in investors who actively compete in financial markets. However, their data is restricted to trades in the Nokia stock, and therefore does not give much insight into flows between assets. Moreover, their network falls short on exhibiting interaction between investors.

2.2 Text Data in Finance

At the time of writing, a series of studies exist that link sentiment, measured through diverse approaches, to stock market performance. Some important ones are reviewed in Gentzkow et al. (2019), but they offer interesting examples that extend beyond sentiment analysis and indeed financial markets. One notable example measures opinions on Twitter using a lexicon approach, finding certain moods to significantly lead changes in the Dow Jones Industrial Average (Bollen et al. 2011).

Such lexicon approaches, whereby specific words are scored based on their prevalence in documents categorised by opinion, are widespread. Further social science research, often studying manifestations of political opinions, use more powerful tools, such as Google's Bidirectional Encoder Representations from Transformers (BERT) algorithm (Devlin et al. 2018). This particular algorithm trains a final layer of nodes from a pre-trained classifier on labelled data. The classifier itself is a neural net, pre-trained by Google on a corpus of Wikipedia entries to i) predict the probability distribution of words appearing in a given sentence (Masked Language Modeling), and ii) predict the relationship between sentences (Next Sentence Prediction).

The study of user interaction and the subsequent expression in opinions is receiving growing attention in these fields. Such features of the data are promising in advancing empirical behavioural finance; as explained in Shiller (2014), 'dumb money' will often rely on opinions of their immediate peers rather than an independent or thorough analysis.

2.3 Opinion Dynamics

The important role that social networks play in forming individuals' opinions and collective behaviours became apparent in the seminal work of Granovetter (1973). More recently, Centola & Macy (2007) ascertain that behaviours spread through 'complex contagion,' requiring multiple exposures, rather than one-shot spreading, termed 'simple contagion' in epidemiology.

The evolution of collective dynamics on social networks has been studied in many contexts, including healthcare outcomes and product adoption, among others (Christakis & Fowler 2008, Lehmann & Ahn 2018). A simple approach, followed in Christakis & Fowler (2008), builds a logistic regression model for the dependent variable at time $t + 1$ as a function of demographic attributes and the status of the dependent variable among contacts at time t. This approach is a fruitful, simple way to gain an estimate of peer influence.

As the field and its applications advanced, causality and the confounding problem of homophily, whereby people that interact share internal, unobserved preferences, emerged as key challenges (Lehmann & Ahn 2018). As a result, Aral et al. (2009) develop a matched sample estimation framework for distinguishing between influence and homophily in adoption of service applications in an instant messaging network. Their compelling conclusion is that previous studies overestimated peer influence by 300-700%.

Given these inherent challenges, progress came through designing controlled studies in social network settings, as in Aral et al. (2009) and Leng et al. (2018). In particular, Leng et al. (2018) offers a promising approach for the discussion network in WSB. The authors build a controlled study by pairing individuals based on observable demographic and social media characteristics from their online presence.

In all, the OD literature has made substantial progress in characterizing the way in which people adopt behaviours from their neighbors across many different applications. Some examples in economics are Banerjee et al. (2013) and Lehmann & Ahn (2018). It offers necessary tools to characterise the spread of interest and sentiment on financial assets through social media.

3 Data

This section presents the rich text data available from Reddit, and motivates the specific attention to the WSB subreddit. Initially, a topic model gives a high-level overview of the dominant narratives that governed users' interests over time. The model detects a mix of conversations, from those that follow specific real world trends, such as the American election or the COVID-19 pandemic, to those that hype positions on a handful of stocks and indices, notably the soaring interest in puts on the S&P 500 index during the COVID-19 crisis. Delving deeper, this section outlines the approach in identifying specific assets from users' exchanges, as well as the sentiments expressed within. Lastly, it presents some preliminary statistical evidence of the forum's close relationship to the market, similar to previous studies on text data in finance detailed in section 2.

3.1 What is *WallStreetBets*?

Reddit, launched in 2005, is a social news aggregation, web content rating, and discussion website which is ranked as the 20$^{\text{th}}$ most visited site globally as of March 2020, with over 330 million anonymous users in 2018[1]. The website contents are self-organized by subject into smaller sub-forums, 'subreddits,' to discuss a unique, central topic.

Within subreddits, users make titled posts, 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 the submission by the amount of upvotes it receives, but lowers it with age. Therefore, the first posts visitors see are i) highly upvoted, and ii) recent. Comments themselves are visible within a post, and are subject to a similar scoring and commenting system.

The WSB[2] subreddit was created January 31, 2012, and reached one million followers in March 2020. As per a Google survey from 2016, the majority of WSB users are 'young, male, students that are inexperienced

[1] `https://en.wikipedia.org/wiki/Reddit`

[2] `https://www.reddit.com/r/wallstreetbets/`

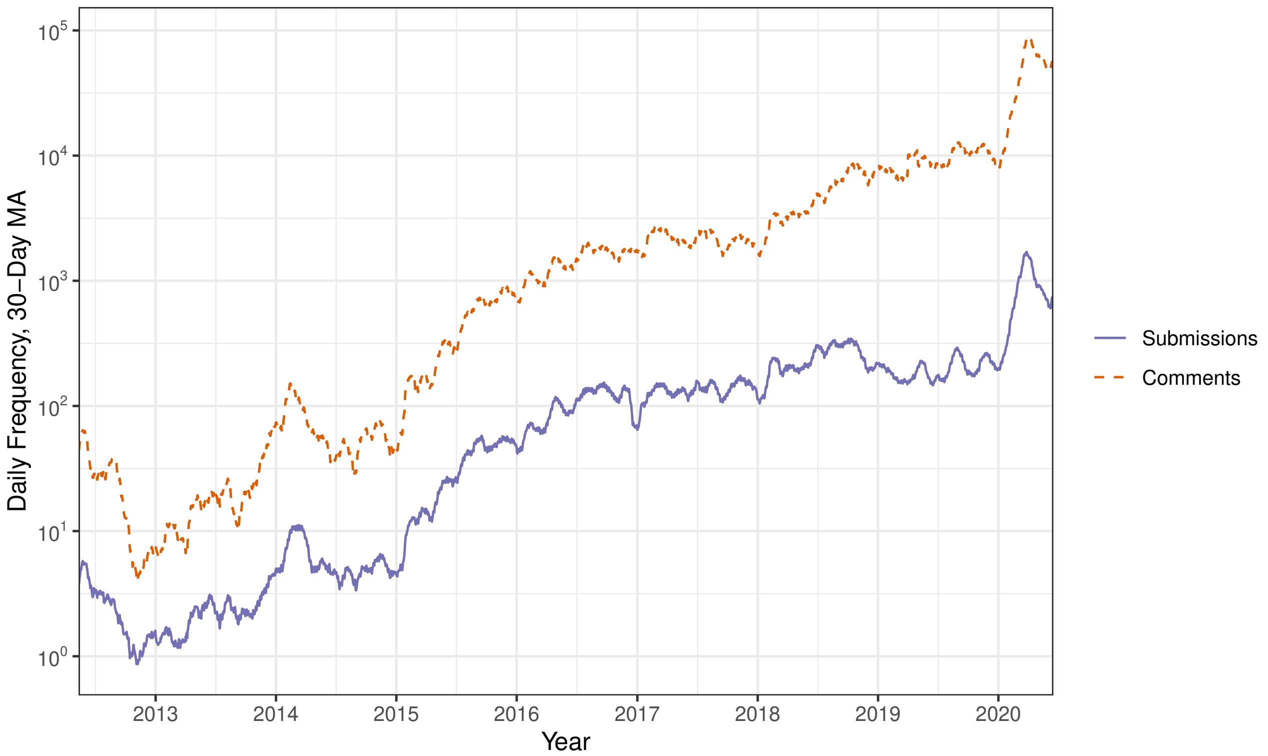


Figure 1: Daily activity on WSB; the daily submission and comment counts, smoothed over 30 days, demonstrate a persistent exponential increase from 2013 to 2020, with a substantial jump in early 2020.

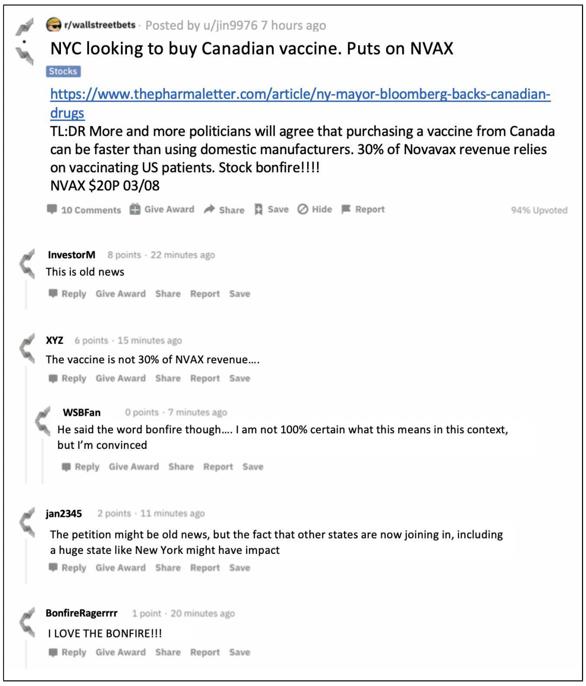
investors utilizing real money (not paper trading); most users have four figures in their trading account'[3]. The conversation guidelines outlined by the moderators handily demonstrate the tone of the conversations:

- Discussion about day trading, stocks, options, futures, and anything market related,
- Charts and Technical Analysis,
- Shower before posting,
- Some irresponsible risk taking,
- People sharing trades, ideas, observations.

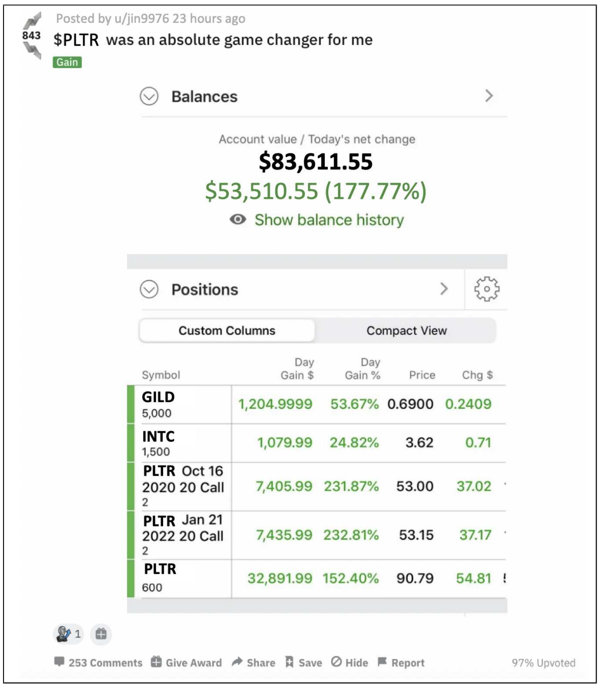
The subreddit's size grew steadily since 2015, but two jumps are notable in Figure 1; a smaller, seemingly idiosyncratic rise in early 2018, and a sharp spike during the COVID-19 pandemic. Figure 2a displays a typical exchange on the WSB forum. Individuals discuss some stock-related news and discuss their sentiments on whether this will affect stock prices. In addition to active market discussions, there is ample evidence of users subsequently pursuing the investment decisions encouraged in these conversations. They post screenshots of their investment gains and losses, which moderators are encouraged to verify, as illustrated in Figure 2b. These observations are reminiscent of Shiller (2005) in defining an asset bubble:

> A situation in which news of price increases spurs investor enthusiasm which spreads by psychological contagion from person to person, in the process amplifying stories that might justify the price

[3] https://andriymulyar.com/blog/how-a-subreddit-made-millions-from-covid19



(a) A Typical Discussion on WSB



(b) Sample Screenshot of User Gains / Losses Posted on the WSB Forum

Figure 2: What Does WSB Look Like? These figures are based on WSB forum discussions and posts however the exact text, usernames, and conversation details have been modified to protect user identities.

> increases and bringing in a larger and larger class of investors, who, despite doubts about the real value of an investment, are drawn to it partly through envy of others' successes and partly through a **gambler's excitement**.

Retail investors are gaining increased power in determining asset prices. The content found on WSB, in addition to surveys, suggest that this is indeed a valuable source of data for understanding how retail investors reach consensus. In particular, the 'casino' quality of the exchanges on the forum offers unprecedented insight into agents' risk preferences.

All posts made on Reddit, plus their metadata, can be queried via Reddit's API, as well as other sources. In what follows, we downloaded posts on WSB using the PushShift API[4]. The only caveat of PushShift is that all data are recorded in at the time of posting. Therefore, real-time meta-data, such as upvotes, downvotes, and comment counts are not updated. We retrieved them by separately querying PRAW, Reddit's direct API.

The full dataset consists of two parts. The first is a total of 452,720 submissions, with their authors, titles, text, timestamp, and upvote scores. The second is comprised of 15.4 million comments, with their authors, text, timestamp, upvote scores, and their linked comment or post. The following sections will predominantly rely on submissions for text data, since they are substantially richer. Comments are largely used to trace user activity and, subsequently, the interaction between discussants.

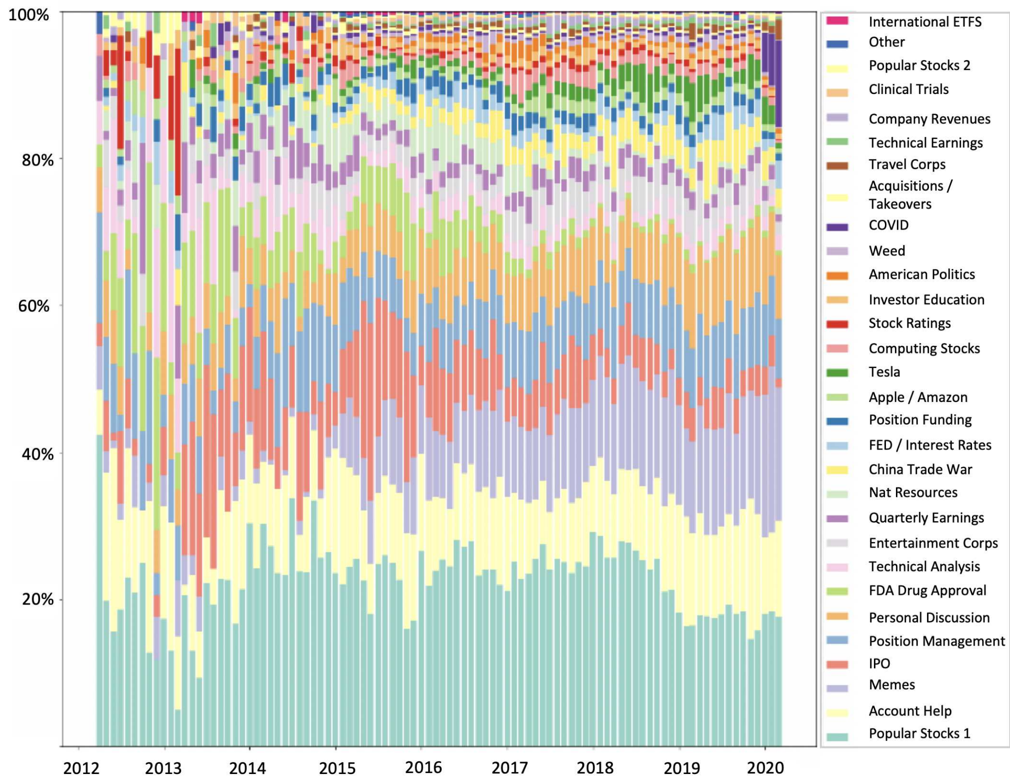


Figure 3: Temporal Trends in Topics

3.2 Dominant Narratives

Does WSB reflect new information for the larger market to trade on, or social activity that drives perceived changes in value, regardless of fundamentals? A causality between sentiment and asset prices is firm, to the extent that the latter dominates. A topic model is an elegant method to evaluate the content of WSB discussions. The results from three different topic models are remarkably robust, as most topics include similar keywords and appear consistently. Superficially, certain topics correlate strongly with market indices despite being thematically different. Typically, generic discussions on mainstream tickers and casual conversations are more prominent one to two months after the stock market produces larger-than usual returns. In contrast, discussions on economic policy prevail in months coinciding and preceding market downturns. Full details are available on request.

Figure 3 presents our preferred topic model, namely the Biterm Topic Model (BTM), which is optimal for smaller bodies of text (Yan et al. 2013). Submissions from April 2012 to April 2020 give a time series of almost 100 months. Figure 3 is a stacked plot of the monthly submission count by each topic, normalised by the total. On one hand, some topics persist in the overall discussion. For example, discussions on a handful of popular stocks and educating users hold a similar share of the overall discussion across all time periods. On the other hand, topics concerned with larger economic trends wax and wane over the observation period. Two examples of this are the uptick in submissions discussing marijuana stocks, coinciding with the drugs' legalisation in select US states, and the COVID-19 topic, which is negligible until January 2020, but dominates in every subsequent month. 'Memes' also grew in influence, likely due to the rise of satirical content entwined with the growing

[4] https://pushshift.io/

popularity of the forum.

3.3 Identifying Tickers

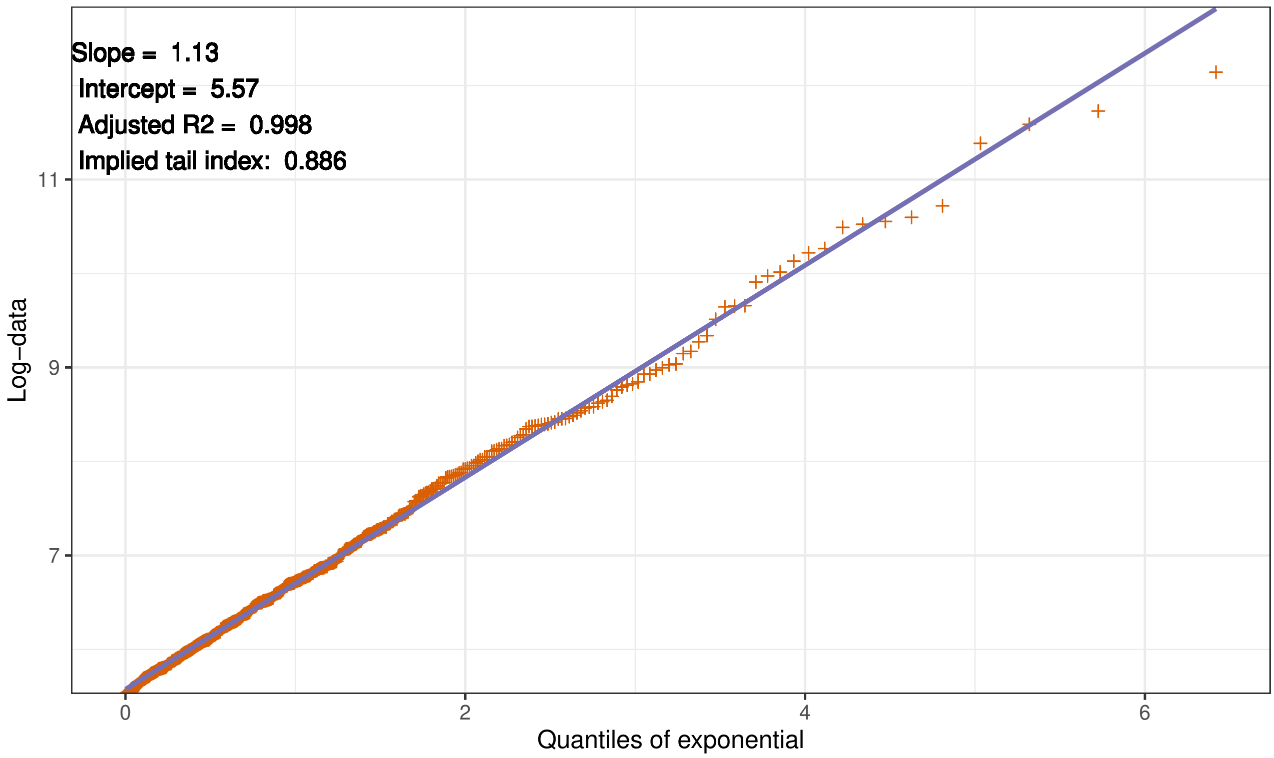


Figure 4: Tail in ticker mentions on WSB; the log of the count an asset is mentioned and commented on is plotted against the theoretical quantiles of an exponential distribution. The linear fit suggests the data exhibit a heavy tail, such that a few assets dominate discussions while the majority are only mentioned sparingly.

In order to understand how users interact around certain assets, we filtered the text data for mentions of specific tickers. 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. A first match is made by identifying any succession of two to five capital letters. Subsequently, we used a pre-determined list, scraped from Yahoo Finance and Compustat, 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 list of such tickers, and ignored featured matches, to build a preliminary list of candidates. 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 mention of '$CEO' or '$a' counts as 'CEO' and 'A', respectively. We checked these extracts again against the scraped list of available tickers.

Table 3 in Appendix A.1 displays the 20 tickers that feature most heavily in Reddit conversations. These are typically stocks of IT firms, such as AMD or FB. A handful of indices are also present, notably the S&P 500 (SPY) and the VIX. As expected, a small share of tickers dominate the discourse. This is further emphasised in the tail of the ticker mentions' frequency distribution, for which Figure 4 displays a QQ-plot. The solid line reflects a density from a Pareto distribution, and the orange crosses refer to the actual densities from the data. We arbitrarily cut from the top 10^{th} percentile, which fitted the Pareto distribution remarkably well. The slow decay implied by an estimated tail exponent of approximately 0.9 would suggest that ticker mentions are heavy-tailed, similar to other vocabulary distributions.

In what follows, we used submissions for which a single ticker was identified, forming a body of 124,438 submissions.

3.4 Measuring Sentiments on WSB

The second key measure to gauge in WSB text is an objective metric for the sentiment expressed by users on the tickers in question. Sentiment analysis is a field that is progressing rapidly, and thus plenty of tools are readily available. We preferred Google's BERT algorithm, a pre-trained neural net that set the standard in NLP sentiment analysis tasks. Work not shown here implements an alternative regression-based approach as a robustness check, but BERT is found to perform better out-of-sample.

Out of the 124,438 available submissions, we randomly selected and labelled 2,774 as either neutral, bearish, or bullish with regards to the author's expressed opinion on the future performance of the detected ticker. We trained BERT on 90% of this data, and used the remaining 10% for validation. The results are satisfactory; BERT correctly classifies two thirds of the data, out-of-sample. This is comparable with the regression's *in-sample* accuracy, which is not shown here.

Table 1: WSB Sentiments and Asset Price Movements

	Dependent variable:		
	R^e_{it}	$\log(V_{it})$	$\log(P^H_{it}/P^L_{it})$
	(1)	(2)	(3)
A^+_{it}	.002*** (.001)		
A^0_{it}	.002*** (.0005)		
A^-_{it}	−.004*** (.001)		
$\widetilde{A}^+_{it}$		.08*** (.005)	.002*** (.0003)
$\widetilde{A}^0_{it}$		.12*** (.004)	.01*** (.0002)
$\widetilde{A}^-_{it}$		−.12*** (.01)	−.01*** (.0004)
Model:	Five Factor	Within-Group	Within-Group

*p<0.1; **p<0.05; ***p<0.01

Notes: This table presents OLS estimators for the correlation between aggregated daily sentiments in WSB and three dependent variables. Superscripts +, 0 and − denote bullish, neutral and bearish sentiments, respectively Column (1) uses the daily five factors, download from Kenneth R. French's data library, although estimates are not shown here. Daily excess returns are calculated as the ratio of subsequent closing prices for assets i at times t and $t-1$, as recorded in Yahoo Finance, minus one plus the risk free rate, as reported in French's data library. Columns (2) and (3) regress the variation in sentiment within asset i and time t on the within asset and time variations in log of trading volume. Daily cross-sectional average log volumes and historical asset average log volumes are subtracted from the variable. The same is done in column (3) for the log-difference in daily intra-day high and low prices for asset i.

A simple exercise serves to verify the validity of using sentiments from WSB to gauge market activity. Define the total sentiment i at time t on asset j, $A_{i,jt}$, by summing the number of submissions expressing opinion i on the future price of asset j. The Fama & French (2015) Five-Factor model should detect any explanatory power to asset price fluctuations, in addition to common asset price ratios, these quantities offer.

Table 1 implements this model using price data downloaded from Yahoo Finance[5] and the Five-Factor time series kindly provided by Kenneth French's data library[6]. We aggregated submissions over t to match the closing times of subsequent trading days, typically 4pm in U.S. markets, to prevent overlap between observed sentiments and future price changes. Column (1) shows the coefficients of the sentiment quantities of interest. These correlate significantly with asset j's daily excess returns, whereby bullish posts relate to a positive return, neutral post to a somewhat lower return, and bearish posts to a negative return.

Columns (2) and (3) present two additional results of interest. Aggregate sentiments correlate significantly with the associate asset's daily trading volume, as well as the intra-day trading range. This indicates that users on WSB are keen to latch onto volatile assets, with prospects of high returns yet high risk. Interestingly, the number of bearish posts correlates with lower relative trading volumes and intra-day price ranges.

These results do not indicate causation. While submissions likely raise attention to these assets, alone they fail to gauge whether these movements follow fundamental shifts in the underlying price, or simply an idiosyncratic rise in the demand for that asset. With this caveat in mind, results not shown here indicate that portfolios built on lagged sentiments not only fail to outperform the market, but indeed produce consistent losses.

4 Self-Organised Bull Runs

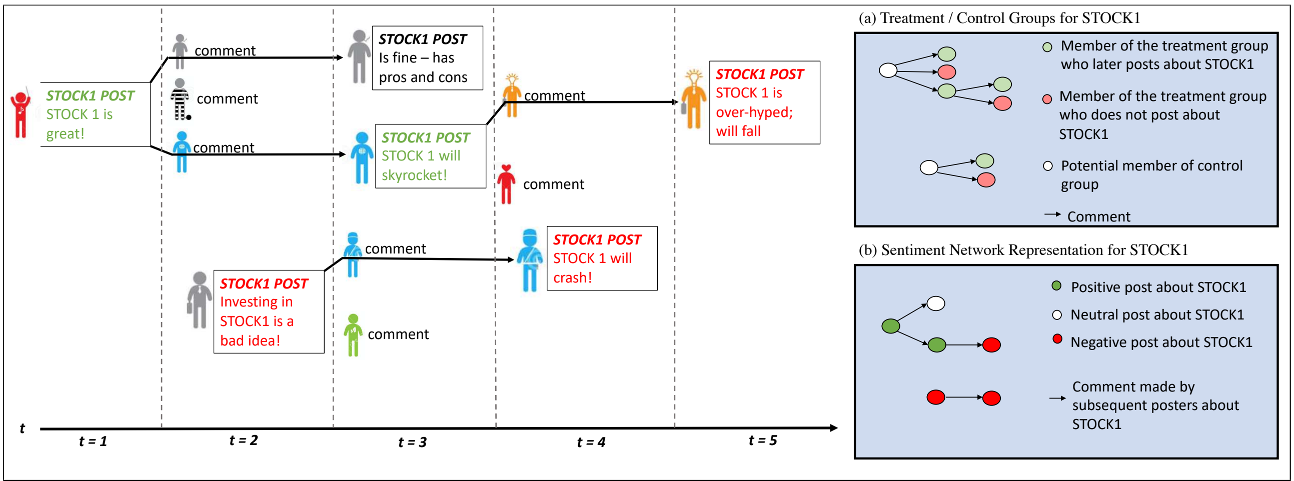


Figure 5: Illustrative Ticker Discussion on WSB

Sentiments correlate with stock market activity, but to what extent do sentiments cause it? Discussions on WSB heavily concentrate around a few tickers, suggesting some form of social contagion, or 'hype.' The main intuition follows the insights of Barberis et al. (2006), but also underscores the communal nature in the discovery of valid gambles that users can pursue. As a result, social interactions amplify the extent to which the nominal price can overshoot the fundamental value after an initial shock.

[5] https://finance.yahoo.com/

[6] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Our overarching goal is to determine the extent to which users' investment interests are self-reinforcing. Simply, *do investors express demand for an asset because of an independently perceived payoff, or because of another user's stated interest*? This type of feedback in a community can be estimated as a spillover rippling through a network of people, linked by their interactions in WSB discourse. Figure 5 displays a toy model for such a network. A user's submission on a ticker with a particular opinion is followed by multiple comments. Some of the commenting users may subsequently post their own submission on the same ticker, and their own opinion on its potential. In this context, contagion in asset demand can be understood as the extent to which submissions are made on a ticker because the author learnt about it from another submission. A second vantage point is to measure the degree to which users will reproduce the sentiments from their peers.

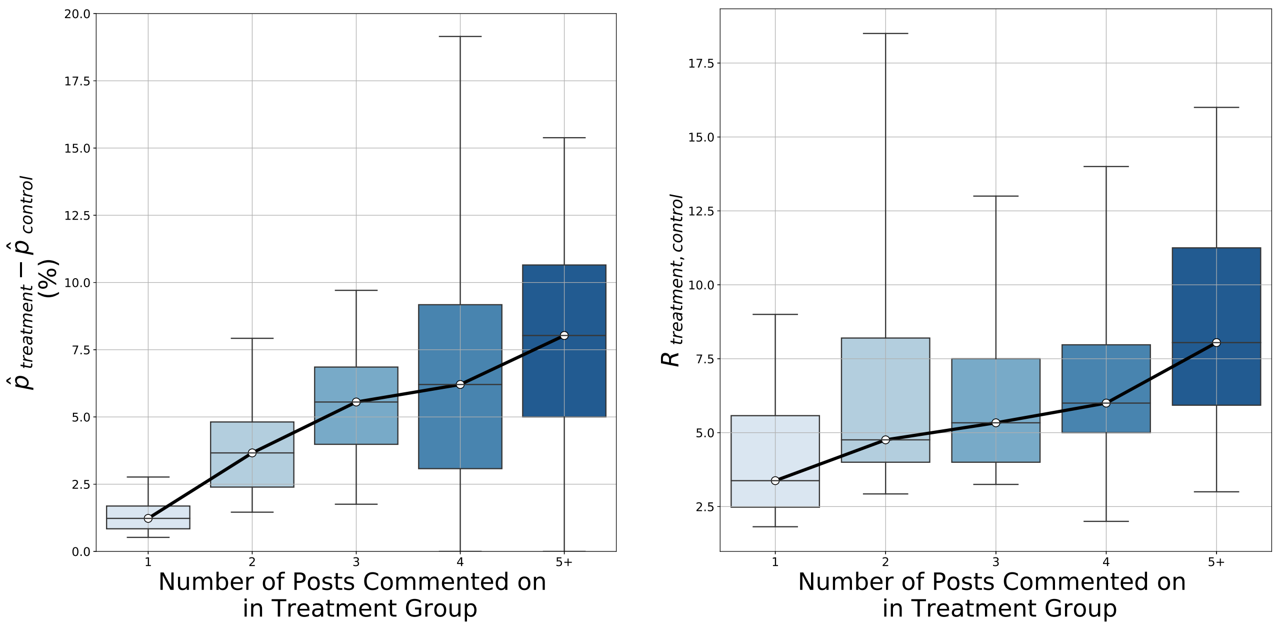
Section 4.1 starts by quantifying the spread in asset demand through the WSB network. This is measured as the propensity for a user to post a submission conditional on them previously commenting on submissions discussing the same ticker, thus demonstrating sustained interest. The main challenge comes in addressing the network's endogenous link formation, as hidden characteristics will determine a user's *choice* to comment. This problem, known as 'homophily' in the broader network literature, leads to an overestimated spillover (Aral et al. 2009). Section 4.1 addresses this issue by matching on observables, a method outlined in Leng et al. (2018), in order to estimate a causal link between a user's engagement with a stock and her future posting activity on that ticker. The method relies on the vast quantities of data available on each user's history on Reddit.

Subsequently, section 4.2 studies the contagion of sentiment: if a user engages in a discussion about an asset with an expressed sentiment, what is the probability she adopts the same sentiment in her own submission? The results demonstrate that the WSB network exhibits significant sentiment contagion, as older submissions influence future sentiments to the degree that users interact.

4.1 Contagion in Asset Demand

Figure 4 demonstrates that discussion sizes of tickers are heavy-tailed. While most tickers are mentioned fewer than five times over seven years, some are mentioned in thousands of submissions, accompanied by tens of thousands of comments. Typical network formation models, such as preferential attachment, predict that these tail tickers emerge from contagion: people prefer to discuss a stock that friends are already debating. Therefore, an accurate estimate for the probability that any one individual posts a new submission about an asset, given previous exposure to discussions on this asset, offers a measurable quantity for contagion in asset demand. For this reason, in addition to computational constraints, we filter the sample of submissions to the top 1% of tickers by number of mentions.

The key challenge is to control for homophily, as similar users create similar submissions regardless of their previous interactions. Specifically, one of the most important forms of homophily to control for is the exposure of individuals to movement of the asset: one may take interest in a stock because it experiences outsized returns, rather than hearing about it from someone else. To tackle this, we matched individuals who comment on a submission with those who did not comment, but were active on the forum shortly after the submission was made. This effectively controls for exposure to the same market moves, as well as associated news. In order to control for overall behavioural patterns, we matched individuals based on similar posting and commenting patterns, and on similar activity times on WSB. Individuals who commented on a tail ticker submission form the *treatment group*, and are matched to users who did not comment, the *control group*. Comparing each group's



(a) Difference in Proportions

(b) Ratio of Proportions

Figure 6: Estimated Effect of Commenting on Posting a new Submission; the distribution of the estimated treatment effect ($\hat{P}_j^T - \hat{P}_j^C$) and the associated odds ratio ($\hat{P}_j^T / \hat{P}_j^C$) across the top 1% of tickers is graphed as a series of box plots. The solid line connects the median estimate, and the protruding dashes represent the 5$^{\text{th}}$ and 95$^{\text{th}}$ percentiles.

propensity to post a submission about the same ticker yields the extra interest stemming from social contagion.

Following complex contagion theory, the transmission of interests and beliefs may require multiple sources of activation, or social reinforcement, to spread (Lehmann & Ahn 2018). In order to quantify the level of social reinforcement necessary to drive a transmission in interest, users were grouped for every ticker by their commenting frequency on submissions related to that ticker. In total, about 90% of users commented on fewer than five different submissions, and we grouped those who commented on five or more submissions together. We matched each group separately to different control groups. Regardless of the number of submissions they commented on, we do not include commenters on the same ticker in any control groups. Henceforth, we considered the number of different submissions that an individual commented on (prior to their submission) the amount of treatment the individual received. We use five treatment groups: people who commented on one, two, three, four, or over five different submissions. These categories match the axes in Figures 6 and 7.

We matched users on i) whether they were active on the WSB forum for approximately the same time period, ii) whether they have similar commenting or posting characteristics in the forum, and iii) whether the member of the control group was active on the forum when the relevant submission is visible. Thus, this exercise matches users not only on observed characteristics, but also on their exposure to news and price moves. Appendix A.3 offers the full details.

The first result of interest is the difference is the observed proportion of submissions in the treatment and control groups, $\hat{P}_j^T - \hat{P}_j^C$, where

$$\hat{P}_j^T = \mathbf{E}(P_{i,j,t_2} = 1 | Q_{i,j,t_1} = 1), \tag{1}$$

$$\hat{P}_j^C = \mathbf{E}(P_{i,j,t_2} = 1 | Q_{i,j,t_1} = 0), \tag{2}$$

and P_{i,j,t_1} is one if user i posts a submission on asset j at some point after commenting, Q_{i,j,t_1} is one if user i previously commented on a submission about asset j and $(t_2 > t_1)$ ensuring that the submission occurs after commenting. Q_{i,j,t_1} is only observed once for each individual i. Therefore, we construct the counterfactual, $\hat{P}_j^C$, from the users matched on observables to the commenters. The difference in proportions measures the likelihood for a user to post a submission given they are treated, and not because of unobserved characteristics.

Figure 6a presents the difference in proportion estimates as a function of the frequency with which a user has commented. The first important result highlighted is that all differences in proportion are positive, implying that engaging in discussions on WSB about an asset strictly increases future interest in the asset. This effect becomes more prominent the more a user comments on ticker-related posts, evidencing some form of threshold contagion in asset demand. However, the small samples for users who commented four or over five times introduce large confidence boundaries.

Therefore, asset demand on WSB is partially endogenously determined, to the extent that it is proxied by the number of submissions made on an asset. People become interested in an asset because others are discussing it, and not necessarily because of its fundamentals. The effect in Figure 6a appears small in absolute terms because there is one order of magnitude more comments than submissions, as seen in Figure 1. However, the large estimated odds ratios in Figure 6b, $\hat{P}_j^T / \hat{P}_j^C$, indicate that commenting leads, on average, to a fourfold increase in the probability of authoring a new submission on the same ticker when an individual is exposed to a single treatment, with the ratio increasing as a user comments on more ticker-related posts.

Figure 7: Bias from Homophily in Estimates for Contagion in WSB; the average difference in the treatment effect – the percentage of individuals who create a submission on a ticker – in the treatment and control groups where control groups are selected randomly (filled circles) or matched on observables (open circles).

Does the matching exercise actually help in controlling for unobserved heterogeneity? Consider a randomly selected control group for $\hat{P}_j^C$, denoted $\hat{P}_j^R$. For random matching, we considered all individuals who have been active on the WSB forum (either commented or written a submission). We repeated the following exercise for every ticker: i) removed any commenters (members of our treatment group) on that ticker from the list of all WSB participants, and ii) matched each member of the treatment group to a random individual from those remaining WSB participants. The estimated treatment effect for both *Random Matching* and *Matching on Observables* is plotted in Figure 7. The result is striking; *Random Matching* overestimates the treatment effect between 140-380%, underscoring the importance of controlling for homophily.

In summary, when comparing similar users on WSB there is clear evidence that they talk about assets, not only because of their fundamentals or users' characteristics, but also because other users talk about them. This lends support to the hypotheses in behavioural finance that emphasise the role of social interaction in driving asset demand and their subsequent price fluctuations. The remaining question, addressed in section 4.2, is whether users reproduce their companions' beliefs on future performance, in addition to attention alone.

4.2 Sentiment Contagion

Section 4.1 establishes the effectiveness of using comments on WSB to determine user demand, or interest, in an asset, given by their probability of discussing it in a separate submission. The question remains whether this community actively forms a consensus on the direction the asset will take in the future to inform their investment decisions. To this end, this section studies the network of user interactions on WSB in relation to the sentiments detected in user submissions.

We formed a network in three steps. First, we extracted all submissions that contain a single, unique ticker; these are the nodes within this network. We place a directed edge, from an earlier submission to a later one, if the author of the later submission commented on the earlier submission. The network construction process is exemplified in Figure 5. Submissions with the same author are not linked, since authors may comment on their own submission. The resulting graph characterises the network, scrutinised in section 4.1, through which sentiments spread. If an individual commented on another's submission about a ticker, prior to expressing her own opinion about this ticker, their opinion may be influenced by the initial submission. This section constructs implements this network to track the flow of sentiment from one submission to the next.

Four examples are displayed in Figure 8. Figures 8a and 8b display the network of submissions that mention a ticker experiencing a considerable price increase: NVDA, which rose by 82.4%, and TSLA, which rose 148.9%. In contrast, Figures 8c and 8d are drawn when the prices of SNAP shrank by 51.6%, that of the SPY ETF by 24.1%. Two patterns stand out. First, the overall sentiments were abundantly bullish in the former two cases, as seen by the extent of blue nodes, as opposed to the overwhelmingly bearish sentiments in the latter, demonstrated by the extent of red nodes. The second pattern is that the magnitude of activity was much more pronounced during the smaller downturn of the S&P 500 (SPY), suggesting that external popularity drives user activity. This is plausible given the amount of media coverage on the economic fallout of the COVID-19 pandemic.

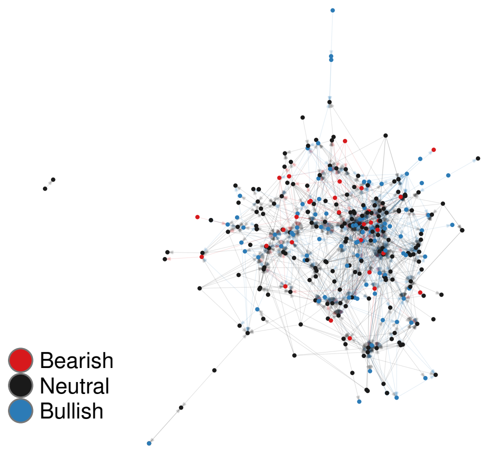
The extent to which the sentiments between linked submissions correlate, net of the actual price movement relationship detailed in section 3.4, would count as a spillover from social interaction. Visually discerning communities of bullish, bearish or neutral posts in Figure 8 is difficult, but these spillovers can be estimated

Table 2: WSB Sentiment Spillovers

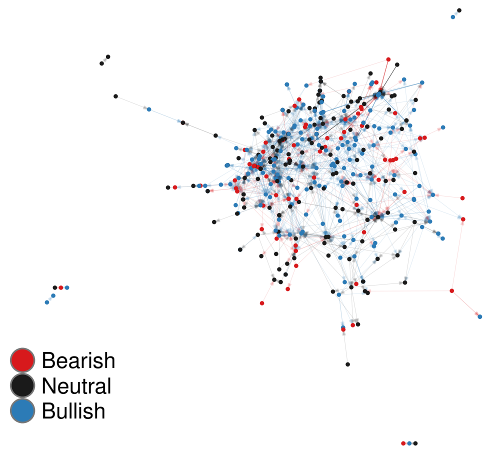
	Dependent variable:							
	a^-	a^+	a^-	a^+	a^-	a^+	a^-	a^+
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
R^e	$-.58^{***}$ (.08)	.02 (.04)	$-.58^{***}$ (.08)	.02 (.04)	$-.60^{***}$ (.08)	.01 (.04)	$-.59^{***}$ (.08)	.01 (.04)
$\widetilde{R^e}$	$-.36^{***}$ (.04)	$.05^{*}$ (.02)	$-.35^{***}$ (.04)	$.05^{*}$ (.02)	$-.33^{***}$ (.04)	$.05^{**}$ (.03)	$-.33^{***}$ (.04)	$.05^{**}$ (.03)
σ	$-.56^{***}$ (.14)	$-.74^{***}$ (.10)	$-.56^{***}$ (.14)	$-.77^{***}$ (.10)	$-.63^{***}$ (.15)	$-.80^{***}$ (.10)	$-.63^{***}$ (.15)	$-.80^{***}$ (.10)
a^+_{-1}	$-.45^{***}$ (.03)	$.46^{***}$ (.02)	$-.44^{***}$ (.03)	$.47^{***}$ (.02)	$-.46^{***}$ (.04)	$.48^{***}$ (.02)	$-.46^{***}$ (.04)	$.49^{***}$ (.03)
a^0_{-1}	$-.83^{***}$ (.03)	$-.44^{***}$ (.02)	$-.82^{***}$ (.03)	$-.42^{***}$ (.02)	$-.84^{***}$ (.03)	$-.44^{***}$ (.02)	$-.83^{***}$ (.03)	$-.43^{***}$ (.02)
a^-_{-1}	$.80^{***}$ (.04)	$-.10^{***}$ (.04)	$.78^{***}$ (.04)	$-.08^{**}$ (.04)	$.82^{***}$ (.04)	$-.08^{**}$ (.04)	$.83^{***}$ (.04)	$-.08^{*}$ (.04)
a^{NA}_{-1}	$-.25^{***}$ (.02)	$-.05^{***}$ (.01)	$-.25^{***}$ (.02)	$-.06^{***}$ (.01)	$-.25^{***}$ (.02)	$-.06^{***}$ (.01)	$-.26^{***}$ (.02)	$-.07^{***}$ (.02)
$\boldsymbol{W a^+}$			$-.08^{*}$ (.04)	$.08^{***}$ (.03)	$-.12^{**}$ (.05)	$.06^{*}$ (.03)	$-.06$ (.08)	.01 (.06)
$\boldsymbol{W a^0}$			$-.16^{***}$ (.04)	$-.22^{***}$ (.03)	$-.20^{***}$ (.04)	$-.23^{***}$ (.03)	$-.26^{***}$ (.08)	$-.27^{***}$ (.05)
$\boldsymbol{W a^-}$			$.50^{***}$ (.05)	$-.10^{**}$ (.05)	$.44^{***}$ (.06)	$-.11^{**}$ (.05)	$.34^{***}$ (.10)	$-.14^{*}$ (.08)
$\boldsymbol{W} R^e$					$.45^{**}$ (.22)	.22 (.14)	$.45^{**}$ (.22)	.22 (.14)
$\boldsymbol{W} \widetilde{R^e}$					$-.54^{***}$ (.14)	$-.10$ (.09)	$-.54^{***}$ (.14)	$-.09$ (.09)
$\boldsymbol{W} \sigma$					$.98^{**}$ (.46)	$.60^{*}$ (.36)	$.98^{**}$ (.46)	$.59^{*}$ (.36)
$a^{\text{NA}}_{-1} \times \boldsymbol{W a^+}$							$-.09$ (.09)	.06 (.07)
$a^{\text{NA}}_{-1} \times \boldsymbol{W a^0}$							.09 (.09)	.06 (.06)
$a^{\text{NA}}_{-1} \times \boldsymbol{W a^-}$							.14 (.12)	.04 (.10)
α	$-.73^{***}$ (.01)	$-.13^{***}$ (.01)	$-.73^{***}$ (.02)	$-.10^{***}$ (.01)	$-.73^{***}$ (.02)	$-.10^{***}$ (.01)	$-.72^{***}$ (.02)	$-.09^{***}$ (.02)
Model	Baseline		Spatial Lag		Spatial Durbin		Spatial Durbin with Interaction	
AIC	186,146.5		185,959.7		181,338.6		181,345	
McFadden's R^2	.0105		.0116		.0362		.0363	
N	92,144		92,144		89,888		89,888	

*p<0.1; **p<0.05; ***p<0.01

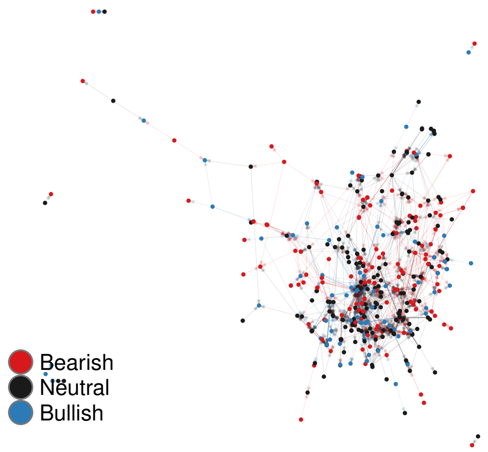
Notes: This table presents estimated log-odds coefficients for multinomial logit models on the expressed sentiment in WSB submissions. Odd columns estimate the log-odds ratio of a sentiment in vector a being bearish (expressed as a dummy of 0 or 1). Similarly, even columns estimate the ratios for a being bullish. Neutral submissions serve as the benchmark.



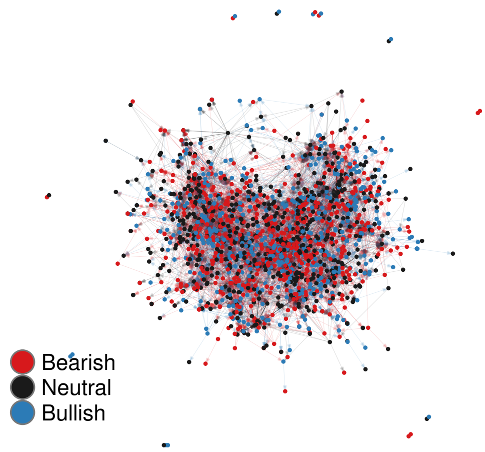
(a) NVDA - April 2017 to December 2017



(b) TSLA - November 2019 to February 2020



(c) SNAP - July 2018 to December 2018



(d) SPY - February 2020 to April 2020

Figure 8: Sentiments in WSB conversations; when the price of an asset increases(decreases) dramatically, individuals post more submissions that are bullish(bearish), with some visual evidence that similar sentiments tend to cluster.

using a network model. Denoting the vector of submissions a with sentiment s, a^s, the following multinomial logit problem isolates the spillover coefficients of interest using a standard spatial lag model:

$$\log\left(\frac{P(a^s = a^s)}{P(a^s = a^0)}\right) = \alpha \boldsymbol{I}_N + \boldsymbol{X}\boldsymbol{\beta} + \sum^{s} \theta_s \boldsymbol{W}\boldsymbol{a^s} + \boldsymbol{\varepsilon}, \tag{3}$$

where the dependent variable is the $1 \times N$ vector of the log probability of recording sentiment $s \in \{+, -\}$ in each submission a^s, over the neutral benchmark $s = 0$. α is a constant, and $\boldsymbol{X}$ includes additional control variables, namely the excess return of the extracted asset on the submission date, R^e, the associated one-week excess return, $\widetilde{R^e}$, and the standard deviation in excess returns over the last five trading days, σ. These are intended to control for the performance of the underlying asset, which, as clarified in Table 1, factor into the sentiment of a given submission. We included an additional set of dummy variables, a^s_{-1}, to identify the sentiment of the same author's previous submission on the same ticker. This checks whether the author's opinion is new ($a^{NA}_{-1} = 1$), the same ($a^s - a^s_{-1} = 0$), or different ($|a^s - a^s_{-1}| = 1$) as their previously revealed sentiment.

These covariates in isolation form the baseline model for user sentiment.

The spatial lag model augments this baseline by introducing $\boldsymbol{W}$, the normalised adjacency matrix for the network described at the start of this section and displayed in Figure 8. Elements in $\boldsymbol{W}$, $w_{i,j}$, equal the number of comments made by i's author on an older submission j with the same extracted ticker, divided by the total number of comments made by i's author on all older submissions with the same extracted ticker. With neutral sentiments as the benchmark case, scalar θ_s is the log-odds of submission a expressing sentiment s given that fraction $w_{i,j} \leq 1$ neighbours express sentiment s. Note that, by taking the sum over all possible sentiments, θ_s also measures the likelihood of dissenting opinions. $\boldsymbol{\varepsilon}$ denotes the error term.

An alternative model of interest is the Spatial Durbin model, which includes spatial lags of covariates in $\boldsymbol{X}$. As far as sentiment transmission goes, this model is useful because it considers the price movements of the discussed asset when the current author commented on the linked submission(s). The extra term's effect is measured by $\boldsymbol{\delta}$ in an augmented form of Eq. 3:

$$\log\left(\frac{P(a^s = a^s)}{P(a^s = a^0)}\right) = \alpha \boldsymbol{I}_N + \boldsymbol{X}\boldsymbol{\beta} + \sum^{s} \theta_s \boldsymbol{W}\boldsymbol{a^s} + \boldsymbol{W}\boldsymbol{X}\boldsymbol{\delta} + \boldsymbol{\varepsilon}. \tag{4}$$

Table 2 presents the results of the multinomial model specified in Eqs. 3-4. For further details, we outline the construction of each variable in Appendix A.2. The baseline model in columns (1) and (2) suggests that submissions are more likely to be bearish versus neutral on days where the excess return of the asset is low. The estimated coefficient of -.58 translates to an increase in the probability of the submission being bearish by almost 6% if the asset price drops by 10% in excess of the market return the day that the submission is made. The converse for bullish posts is smaller, but statistically insignificant. The returns over a trading week is a significant variable; a -10% accumulated return over the preceding five trading days increases the probability of a submission being bearish by 3.7%, and reduces the probability of a submission being bullish by .5%. Moreover, the standard deviation in daily returns observed in the previous five trading days increases the likelihood that a submission expresses neutral opinions. Overall, sentiments follow current and recent changes in the asset prices, but high volatility increases users' expressed uncertainty on the future trajectory of the asset.

Further to price performance, author's sentiments persist over time. Submissions are 58% more likely to express bullishness if they previously made a bullish post, all else equal. The equivalent effect for bearishness is more significant; a submission is more than twice as likely to be bearish if its author posted a bearish submission in the past. However, authors who post a submission for the first time, or expressed equivocal sentiments in their previous post, are significantly more likely to express a neutral opinion.

These estimates for the baseline do not vary substantially across the various specifications, and so will not be further discussed. We present the remaining coefficients estimated for Eq. 3 in columns (3) and (4). Departing from the baseline model, the data demonstrates that authors who previously commented on a bearish post in isolation are 65.8% more likely to express bearish over neutral sentiments, and 9.5% less likely to express bullish sentiments over neutral sentiments. Similarly, but less markedly, authors who previously commented on a single bullish submission are 8.6% more likely to write a bullish submission, yet 7.4% less likely to write a bearish one. Comparable results are also observed for neutral posts.

The results offer strong evidence of the endogenous spread of sentiments between users and discussions. Interestingly, the effect is more pronounced for assets with bearish outlooks. One hypothesis is that users rely on discussions timing a potentially lucrative downturn in the price of an asset. This is in part supported by the spatial Durbin model detailed in Eq. 4. The estimated coefficients in columns (5) and (6) imply that a submission is 4.6% more likely to be bearish if the author commented on a submission made on a day the asset's price increased by 10%. In addition, the volatility at time of commenting being higher increases the probability of bearishness; a 10 percentage point increase in the 5-day average deviation of excess returns from the mean translates to a 10.3% higher probability of bearishness. The reverse does not hold for bullishness.

The final set of columns in Table 2 investigate whether sentiments detected in users who post their first submission versus those that have already expressed an opinion in the past are noticeably different. The answer is in the negative; interacting the relevant dummy variable, a_{-1}^{NA}, with the spillover terms does not yield any statistically significant results.

As a whole, section 4 presents strong evidence that WSB investors form consensus on assets, upon which they actively trade. This consensus is achieved faster for bearish cases than bullish ones, net of actual asset returns and volatility. Combined with the finding in section 4.1 that users converge on heavily-discussed assets for these trades, the observed behaviours follow the hypothesis that large fluctuations in these asset prices are indeed driven in part by feedback loops in those, or possibly similar, investor communities.

5 Conclusion and Further Goals

This paper statistically characterizes psychological contagion in the WSB investor forum. First, section 4.1 provides evidence supporting the fact that investor interest, as expressed in a WSB posts, has a statistically significant endogenous component. The top 1% most discussed assets are estimated to reproduce at four times the rate for users involved in existing discussions as compared to their unexposed counterparts. Users also impact each others' sentiments: the sentiment expressed in a post influences the future sentiments expressed by others, as demonstrated in Section 4.2. This sheds light on how individuals reach consensus on investment decisions, depending on which assets captures interest and the expected trajectory of its price. Three outstanding questions will help synthesise the broader implications of these results.

First, can behavioural finance inform on a theory for the observed behaviours? The appeal of risky gambles and social interaction is established in the literature reviewed as part of section 2. The two models would not be hard to reconcile, and would help to conceptualised the degree of narrow framing users on WSB are subject to, conditional on traditional measures of risk aversion. The network approach is also well grounded in an existing literature that pushes preferential attachment as a driver for the observed behaviours. The particular phenomenon we seek to shed light on is the strong influence of bearish sentiment relative to bullish.

Second, the user interaction network offers rich information in its evolution over the years, for a cross-section of assets. One simple metric that can be generalised from this is the number of feedback loops these discussions exhibit over time. Centrality in the network offers additional information on the prevalence of leaders who sway large parts of the conversations. Generalising, we can model a transition matrix for a markov chain; each asset follows nine states of sentiment and asset returns pair, $[a^s_{(t-1)}, R^e_t]$ where a^s takes on values (-1,0,1) for (bullish, neutral, bearish) sentiment and R^e takes on values (-1,0,1) for (under-performs, average, out-performs) average market returns.

Finally, the distance matching in section 4.1 should be driven by the data, and not fixed arbitrarily. The vast quantity of user data can be used to calibrate a neural network that estimates the likelihood of any two observed users being the same. Furthermore, an adversarial network built on this trained classifier would give a consistent estimate for the counterfactual likelihood of a user posting on a ticker given their underlying characteristics.

References

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.

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

Barberis, N. & Huang, M. (2009), 'Preferences with frames: A new utility specification that allows for the framing of risks', *Journal of Economic Dynamics and Control* **33**(8), 1555 – 1576.

Barberis, N., Huang, M. & Thaler, R. H. (2006), 'Individual preferences, monetary gambles, and stock market participation: A case for narrow framing', *American Economic Review* **96**(4), 1069–1090.

Bollen, J., Mao, H. & Zeng, X. (2011), 'Twitter mood predicts the stock market', *Journal of Computational Science* **2**(1), 1 – 8.

Centola, D. & Macy, M. (2007), 'Complex contagions and the weakness of long ties', *American Journal of Sociology* **113**(3), 702–734.

Christakis, N. A. & Fowler, J. H. (2008), 'The collective dynamics of smoking in a large social network', *New England journal of medicine* **358**(21), 2249–2258.

Cowles, A. (1933), 'Can stock market forecasters forecast?', *Econometrica* **1**(3), 309–324.

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* .

Fama, E. F. & French, K. R. (2015), 'A five-factor asset pricing model', *Journal of Financial Economics* **116**(1), 1 – 22.

Farmer, J. D. (2002), 'Market force, ecology and evolution', *Industrial and Corporate Change* **11**(5), 895–953.

Gentzkow, M., Kelly, B. & Taddy, M. (2019), 'Text as data', *Journal of Economic Literature* **57**(3), 535–74.

Granovetter, M. S. (1973), 'The strength of weak ties', *American Journal of Sociology* **78**(6), 1360–1380.

Lehmann, S. & Ahn, Y.-Y. (2018), *Complex spreading phenomena in social systems*, Springer.

Leng, Y., Dong, X., Moro, E. & Pentland, A. S. (2018), *The Rippling Effect of Social Influence via Phone Communication Network*, Springer International Publishing, Cham, pp. 323–333.

Manela, A. & Moreira, A. (2017), 'News implied volatility and disaster concerns', *Journal of Financial Economics* **123**(1), 137 – 162.

Musciotto, F., Marotta, L., Piilo, J. & Mantegna, R. N. (2018), 'Long-term ecology of investors in a financial market', *Palgrave Communications* **4**(92), 2055 – 1045.

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.

Yan, X., Guo, J., Lan, Y. & Cheng, X. (2013), A biterm topic model for short texts, *in* 'Proceedings of the 22nd International Conference on World Wide Web', WWW '13, Association for Computing Machinery, New York, NY, USA, p. 1445–1456.

A Appendix

A.1 Most Frequent Tickers on WSB

Ticker	Name	Comment Count	Submission Count	Sum
SPY	S&P 500 Index	178,228	9,356	187,584
AMD	Advanced Micro Devices, Inc.	118,288	5,746	124,034
TSLA	Tesla, Inc.	100,713	6,947	107,660
MU	Micron Technology, Inc.	84,051	3,891	87,942
AAPL	Apple Inc.	43,273	1,952	45,225
SNAP	Snap Inc.	37,990	2,039	40,029
AMZN	Amazon.com, Inc.	36,658	1,652	38,310
MSFT	Microsoft Corporation	35,265	1,926	37,191
NVDA	NVIDIA Corporation	34,389	1,576	35,965
VIX	CBOE Volatility Index	27,716	988	28,704
SPCE	Virgin Galactic Holdings, Inc.	25,758	1,683	27,441
FB	Facebook, Inc.	23,638	1,475	25,113
BYND	Beyond Meat, Inc.	21,451	906	22,357
DIS	The Walt Disney Company	20,364	1,093	21,457
NFLX	Netflix, Inc.	19,154	964	20,118
GE	General Electric Company	14,695	941	15,636
JNUG	Direxion Daily Jr Gld Mnrs Bull 3X ETF	14,501	1,078	15,579
RAD	Rite Aid Corporation	14,613	830	15,443
SQ	Square, Inc.	12,682	839	13,521
ATVI	Activision Blizzard, Inc.	10,622	744	11,366

Table 3: Most Frequent Tickers in WSB

A.2 Data Appendix

Besides the scraped text data from Reddit, we downloaded historical price series on the extracted tickers from Yahoo Finance. The time period, denoted by t, is defined as the 24 hours between closing times on the New York Stock Exchange, namely 4pm EST. For example, a submission made on Tuesday at 5pm EST was categorised in the same time frame as a submission made on the following Wednesday at 10am. This is to mitigate any influence news occurring outside market hours might have on conversations in WSB. The choice of timing may not be consistent with non-US stock markets, but, given that almost all discussed assets are US stocks and indices, this does not affect the results.

We calculated the variables used in Table 2 as follows:

- $R^e_{i,t} = \frac{P_{i,t}}{P_{i,t-1}} - (1 + R^f_t)$, where $P_{i,t}$ is the closing price of asset i on day t, and R^f_t is the daily risk free rate on day t,
- $\widetilde{R^e_{i,t}} = \frac{P_{i,t-1}}{P_{i,t-6}} - (1 + R^f_{t-6})^5$, where the 5-period lag is based on a 5-day trading week,

- $\sigma = \sqrt{\frac{1}{5}\sum_{k=1}^{6}(R^e_{i,t-k} - \overline{R^e}_{i,t-1})^2}$, where $\overline{R^e}_{i,t}$ denotes the average excess return observed between days t and $t-5$,

- $\boldsymbol{W}$ is the normalised submission adjacency matrix, where $w_{i,j}$ is equal to the number of comments made by i's author on an older submission j with the same extracted ticker, divided by the total number of comments made by i's author on older submissions with the same extracted ticker,

- a^s denotes a dummy variable for a submission with sentiment s, where s can be bullish ($s = +$), bearish ($s = -$) or neutral ($s = 0$),

- a^s_{-1} denotes a dummy, as above, for the sentiment of the last post made by the author on the same extracted ticker, where a^{NA}_{-1} represents the lack of such a submission.

A.3 Matching on Observables

We matched individuals in the treatment group matched against people in the control group using the following criteria:

1. We matched individuals from the treatment group to all Redditors active (who have either posted or commented) on WSB (excluding those in the treatment group) who became active on the forum during the same month

2. We filtered out individuals in the potential control group whose last activity (post or comment) on the forum occurred before the submission that was commented on by their match in the treatment group (filtering out individuals who have become inactive before our period of interest finishes)

3. Lastly, we filtered out individuals in the potential control group who have posted about a ticker before the time of the relevant submission. This indicates that the individual is *already* interested in the ticker and the events surrounding our submission of interest would have no effect.

4. We then calculated the behavioural distance between a ticker commenter and the remaining potential matched control group using several distance metrics:

D_1: the time between the submission the commenter interacted with and the next activity of the member of the control group. One of the main sources of homophily in this study is the appearance of news or market moves by a particular ticker. Section 3.4 shows a strong correlation between asset returns sentiment and posting volumes. We controlled for this by making sure that a member of the control group was active shortly after the time of the relevant submission in order to have been exposed to the same market and news environment at the time of their activity.

$$D_1 = min_{t_{i,j} > s_k}[t_{i,j} - s_k] \tag{5}$$

where $t_{i,k} \in$ times of activity of j (a member of control group), s_k is the time of the submission a member of the treatment group, k, commented on.

D_2: the average comment and submission length

D_3: the average comment and submission amount

D_4: the time of day of commenting and posting activity - proxy for demographic information. If individuals are most active on social media at the same time of day, this may be an indication of them being located in similar time zones.

D_5: the standard deviation in the time of day of commenting and posting activity - proxy for demographic information. Where D_4 strives to capture the average time of activity, this parameters captures how consistently individuals behave and whether they might be a frequent individual on social media who logs on at different points in the day.

5. We normalized all distance variables D_i between [0,1] using min-max normalization, summed them ($D = D_1 + D_2 + D_3 + D_4 + D_5$) and computed their inverse (1/D) to use this as our final metric $M_{i,j}$ between a member of the control, i, and member of the treatment group, j. A small M would imply that two individuals are a poor match, a large M implies that they are a good match.

   If the individual comments on multiple submissions, the D_1 metric was computed separately for each submission the individual comments on in order to control for potential exposure to each market event. For individuals who commented on over five different submissions, we considered only the most recent five submissions for the purposes of the distance calculation, in order to proxy recent activity and market events which may have prompted asset interest. For someone who comments on four submissions, the distance calculation between them and a member of the control group would be computed as: $D = D_{1_1} + D_{1_2} + D_{1_3} + D_{1_4} + D_2 + D_3 + D_4 + D_5$.

6. Lastly, we solved the following maximal-matching optimization problem in order to match members of the treatment and control groups:

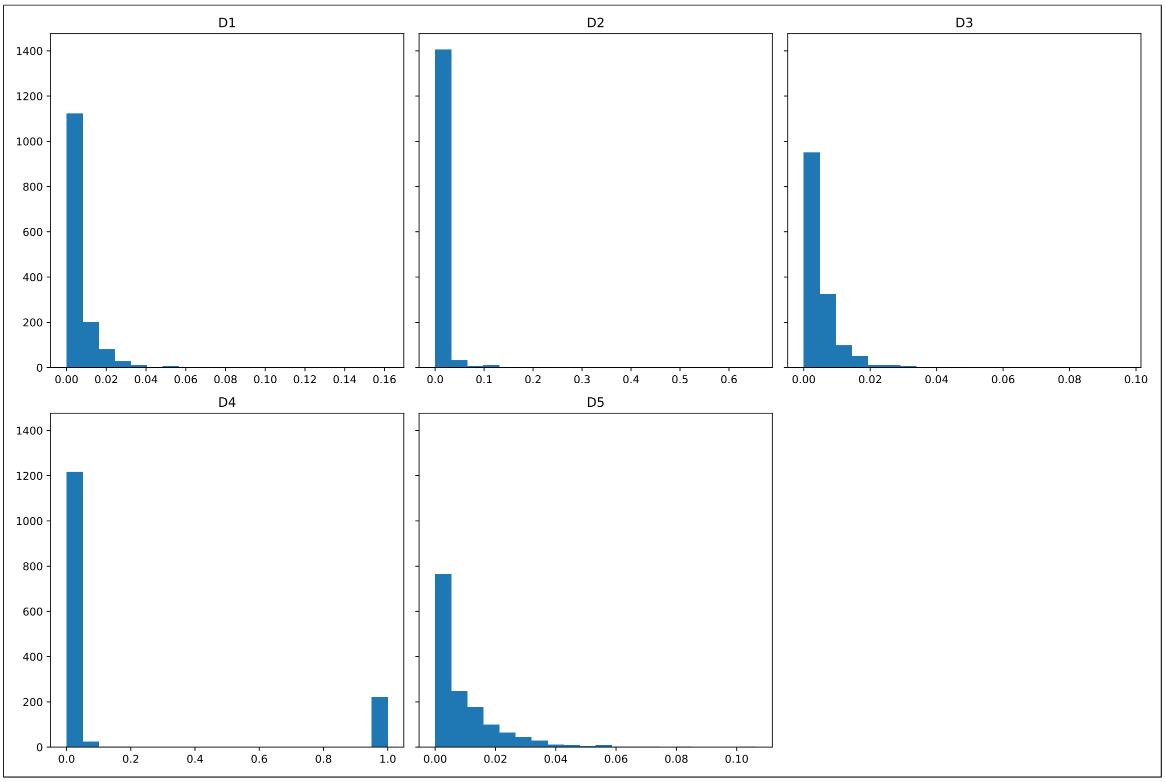
$$\begin{aligned} \max \sum_{i,j \in NXT} & M_{i,j} x_{i,j}, \\ s.t. \sum_{j} & x_{i,j} \leq 1 \;\; \forall i \in N, \\ \sum_{i} & x_{i,j} \leq 1 \;\; \forall j \in T, \\ & 0 \leq x_{i,j} \leq 1 \; \forall N, T, \\ & x_{i,j} \in \mathbf{Z} \; \forall N, T, \end{aligned} \tag{6}$$

   where N stands for our control group, T stands for our treatment group.

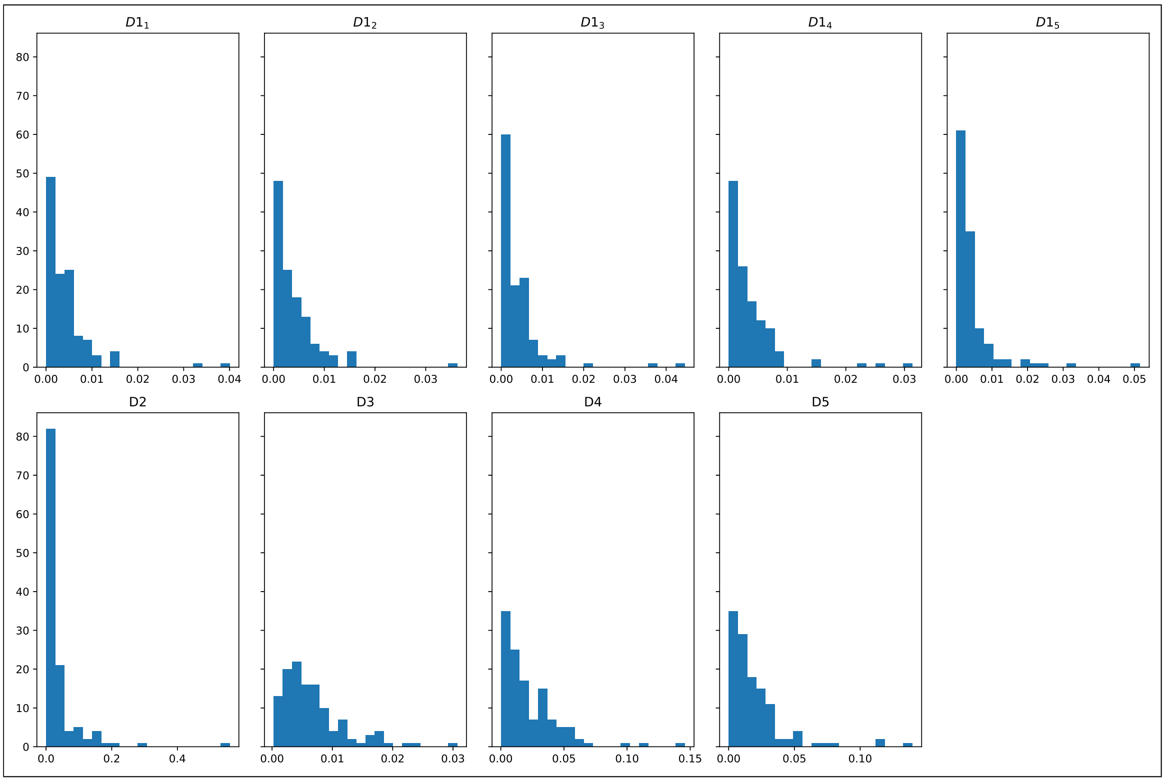
   Practically, even though this is an integer program, we can drop the last integrality constraint.

A question that might arise is what the distribution in specific distance values $(D_1, ...D_5)$ is, and whether, even though over distance D is minimized, the individual distances can take on large values. We examined closely HMNY – the 20$^{\text{th}}$ most popular ticker on WallStreetBets – for the group who only commented once (were exposed to one treatment) and for those who commented five or more times. Figures 9a and 9b display the max-min normalized distributions of the individual distances for the people who comment once and people who comment five or more times on HMNY related posts, respectively. The individual distances seem clustered around 0 implying that, even though in this framework it is theoretically possible for individuals to have large behavioral distances across single parameters, this did not occur in practice. There is a suspicious spike around

1 in Figure 9a, D4, implying that individuals are matched who are maximally distant along this parameter. In practice, individuals who both have no activity before the submission in question will have a maximum value for this distance metric (D4) and 0 for all others except D1. As expected, this spike disappeared for people who commented on five or more ticker-related posts, as they are matched to other who are active.



(a) Distributions of Distances where Individuals in the Treatment Group Comment Once



(b) Distributions of Distances where Individuals in the Treatment Group Comment Five or More Times

Figure 9: Distributions of Normalized Behavioral and Exposure Distances ($D_1, ..., D_5$ detailed above) between Matched Individuals in the Treatment and Control Groups for HMNY.