Conversations in the Market: Sentiment Contagion among Investors*

Haoyang Liu†
liuhy@berkeley.edu
Pengfei Sui‡
psui@cuhk.edu.cn
Federal Reserve Bank of New York
and
The Chinese University of Hong Kong, Shenzhen

Date: February 13, 2022

Abstract

Using a unique data set, we document direct evidence for the role of social interactions on sentiment contagion: investors update their sentiment to more closely match others' attitudes in conversations. Naive investors and male investors are more affected by conversations. Conversations with sophisticated investors, with a high number of posts and with consistent attitudes tend to be more influential. We further document that sentiment contagion predicts investors' trading decisions at the individual level, and trading volume and return volatility at the aggregate market level. Moreover, we confirm the self-enhancing transmission bias in our sample and use it to connect sentiment contagion to the high trading volume during bubble episodes.

Keywords: Social Finance, Sentiment Contagion, Self-enhancing Transmission Bias, Bubbles, Bitcoin
JEL Code:

---

*The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York, or the Federal Reserve System.

†liuhy@berkeley.edu

‡psui@cuhk.edu.cn

1 Introduction

Investor sentiment spans the history of modern finance. For decades, measuring investor sentiment and understanding its effect on markets has been challenging for financial researchers. In past decades, there has been significant progress in studying investor sentiment at the aggregate market level. [1] These works construct various sentiment measurements from available indicators to explain important phenomena in the market. But how does investor sentiment develop at a micro level? Although much work covers aggregate investor sentiment, little is known about the underlying micro dynamics for these aggregate indicators.

In this paper, we fill this gap by providing insights into one micro-foundation for the evolution of investor sentiment. Our emphasis is on investor sentiment contagion, or equivalently how individual investors' sentiments spread from one to another. Such sentiment contagion pattern appears in many important narratives in finance. For instance, as Shiller put it in his seminal book *Irrational Exuberance*, the spread of enthusiasm in markets may have served as the precipitating factor for market booms. Investor sentiment contagion therefore offers a new perspective to understand market dynamics during bubble episodes. Moreover, consistent with the spirit of social finance (Hirshleifer (2020)), we focus on the role of social interactions in the process of sentiment contagion. Therefore, our results also help understand the role of social interactions in affecting market dynamics through shaping and spreading investor sentiment.

Based on a unique data set from an online investor forum, we directly observe conversations between investors and measure sentiment in their conversations. We use the term *Whisper* to refer to other investors' messages in one conversation and focus on the *Whisper sentiment* in our analysis. We document that investor sentiment is positively influenced by Whisper sentiment, an indication that investor sentiment spreads after conversations. Moreover, we find that investor sentiment in our data is not a sideshow: sentiment contagion predicts individual investors' trading decisions, and the intensity of sentiment contagion predicts future return volatility and trading volume at the aggregate market level. Furthermore, consistent with Han et al. (2022), we document a strong self-enhancing transmission bias: investors tend to publish optimistic opinions more frequently after experiencing good returns than bad returns. Self-enhancing transmission bias explains the pervasive optimism in bubbles and indicates a greater intensity of sentiment contagion during bubble episodes. The elevated sentiment contagion explains the high trading volume in bubbles.

The platform we focus on is an online investor community called Bitcointalk. There are several advantages of using this platform for our study. First of all, the platform is created by Satoshi Nakamoto, the founder of Bitcoin, to serve as the official forum for Bitcoin investors to share their opinions and information about the Bitcoin market. The special role of this website in the development of Bitcoin attracts all types of Bitcoin investors to actively participate in it, improving the representativeness of our data. Second, on

[1] Among others, existing literature has proposed market-based measures (Baker and Wurgler (2006)), survey-based measures (Brown and Cliff (2005), Lemmon and Portniaguina (2006), Qiu and Welch (2004)), search-based measures (Da et al. (2014)) and media-based measures (Tetlock (2007), Soo (2018)).

Bitcointalk, investors share their opinions and interact with others by publishing posts and replying to each others' posts. We use textual analysis algorithms to extract investor sentiment from each post. We thus are able to observe investors' sentiment both before and after each conversation and the overall sentiment in the conversation, which serves as an ideal setting to study sentiment contagion. Moreover, a fraction of the investors in this community voluntarily reveals their Bitcoin addresses to enhance the security of their forum accounts. For these investors, we can link sentiment revealed in their posts to their transactions. These unique features help us draw novel conclusions on the role of social interactions through sentiment contagion on market dynamics.

We begin by documenting sentiment contagion. We calculate investors' sentiment change before and after participating in a conversation and regress it on Whisper sentiment in that conversation. We find a positive and statistically significant relationship: as Whisper sentiment increases, investors become more optimistic after conversations. To identify the role of social interactions, we further include investor-fixed effects and date-fixed effects to separate the role of social interaction from variations at the investor level and over time. Controlling for these fixed effects, we find that the Whisper sentiment significantly influences investor sentiment. We also confirm that the effect of Whisper sentiment we identify is not driven by alternative stories such as lagged responses to news or investor attention. Thus, our results highlight the role of social interactions on sentiment contagion.

Investors with different features react differently to Whisper sentiment. First, we document the role of investor sophistication. Using the labels assigned by the Bitcointalk community, we categorize investors as sophisticated investors and naive investors. We find that sophisticated investors are less influenced by the Whisper sentiment. In strong contrast, naive investors tend to hastily update their sentiment: after participating in the same conversation, naive investors are two times more likely to change their sentiment towards the Whisper sentiment. Second, we document a heterogeneous impact related to gender. Compared to male investors, female investors are 9.3% less likely influenced by Whisper sentiment, echoing the previous finding that males tend to be more overconfident (Lundeberg et al. (1994), Barber and Odean (2001)).

We also study which types of conversations are more influential and three features stand out. First, if a conversation contains at least one sophisticated investor, then all else being identical, investors are 5.5% more likely to be affected. One plausible interpretation is that investors treat Whisper sentiment in such conversations as more credible. Second, conversations with more posts tend to be more influential. Bayesian learning can serve as a candidate explanation because investors observe more information from the increased number of posts. Alternatively, a prevailing pattern in psychology literature called the "illusory truth effect" indicates that individuals tend to believe false information to be correct after repeated exposure to that information. Third, more dispersed attitudes in a conversation reduce sentiment contagion, even conditional on the same level of Whisper sentiment. Increased attitude dispersion may make investors less confident in the information they receive in conversation, and therefore would be less motivated to revise their sentiment after participating in the conversation.

Going one step further, we study how sentiment contagion affects investors' trading decisions. We link investors' sentiment change after a conversation to their subsequent trading decisions within a short window of 12 hours. We document that investors' sentiment changes after a conversation predicts the occurrence of their first subsequent transaction and the direction of the transaction. If investors revise their sentiment from neutral to optimistic, then she is 0.202% to buy Bitcoins in the subsequent 12-hour window after the conversation. This finding is robust with or without additional controls for news shocks, market fluctuations, and sentiment on the Bitcointalk community. At a broad level, these results confirm that investor sentiment plays an important role in trading behavior, suggesting that social interactions may have real impacts on the aggregate market.

We then investigate the aggregate impact of sentiment contagion on market dynamics. Specifically, we construct a variable at the community level to trace the intensity of sentiment contagion in the market. We count the number of investors on Bitcointalk who change their sentiment towards the Whisper sentiment in conversations and (after removing the trend) use it as our main proxy for the intensity of social interactions. Since changes in investor sentiment are linked to individuals' future trading behaviors, we conjecture that this variable should at least partially reflect the aggregate demand for position changes in the future, and therefore should predict market dynamics such as the trading volume. Consistent with our conjecture, we document a strong predictive power for both trading volume and future return volatility in the market.

Finally, we shed light on bubbles from the perspective of sentiment contagion. As a starting point, we identify bubble episodes by searching for local peaks in price when there are at least a 100% increase in Bitcoin returns in the past one month and a 40% decline in the subsequent one month. The choice of 100% return aims to conform to Fama's notion that a bubble if it exists, begins with a large price run-up. We find that our choice of threshold can pick most episodes that anecdotal evidence has indicated as bubbles, such as the period around the end of 2017. In these identified episodes, we observe market features that occur frequently in popular bubble narratives, which help justify our approach of identifying bubbles. Specifically, compared to the non-bubble episodes, the daily Bitcoin return is more than 8 times higher and the total dollar trading volume grows by 2.5 times. The return volatility also increases by 55.56%. The sentiment reflected in the news report and the Google Search volume all surges significantly. In summary, our identified bubble episodes feature rapid rises in returns, volume, and market fluctuations, along with increasingly pervasive optimism in media coverage and explosive investor attention.

Shifts in market conditions naturally influence the features of conversations. We document that, during bubble episodes, more investors participate in conversations, investors are more optimistic and the dispersion in attitudes among investors decreases. The explanation we provide for such features is the self-enhancing transmission bias, first proposed in Han et al. (2022), that investors tend to post more frequently after experiencing good returns. We also document that investors become more optimistic after experiencing good returns. Each of these features makes conversations more influential and therefore indicates an elevated intensity of sentiment contagion during bubble episodes.

We use the elevated intensity of sentiment contagion to explain the high trading volume during bubbles. This is a direct extension of our previous findings. We have documented that sentiment contagion in social interactions leads investors to trade and predict the trading volume at the aggregate market level. Since sentiment contagion becomes more intensive during bubble episodes, as a result, we observe the high trading volume in bubble episodes.

Our paper connects to several strands of literature. First, there is a growing interest in the social interaction literature on how investment ideas are transmitted. Shiller (1989) argues that investment in assets is a social activity and Shiller and Pound (1989) started to consider the role of social interaction on the transmission of financial information. Han et al. (2022) is one of the pioneer papers that offer theoretical social approaches to understanding investment decisions and asset prices. Consistent with their analysis, our paper provides direct evidence for the effect of social interactions on investor beliefs and demonstrates that social interactions have real effects on market dynamics by propagating the spread of investor sentiment.

Our paper is very related to Huang et al. (2021). Using stock-financed M&A as an exogenous shock, they document the contagion of abnormal trading activity from "infected" investors to their neighboring investors. Their findings are consistent with the mechanism in which investors communicate with their local peers on trading decisions. They also estimate the rate of communication among different investors. Similar to their results, we also highlight the concept of contagion, but our emphasis is on investor sentiment and the impact. We document novel connections between sentiment contagion and the subsequent trading decisions at the individual level. We also rely on the self-enhancing transmission bias to link sentiment contagion to high trading volume during bubble episodes.

Our paper is closely related to studies that analyze the influence of peer actions. Hong et al. (2004) and Brown et al. (2008) provide evidence consistent with the notion that individuals are more likely to participate in the stock market when their geographically proximate peers participate. Hong et al. (2005) also shows that investors tend to buy stocks their local peers have been buying in the recent past. However, most of these studies compromise by relying on indirect measurements, such as distance, to proxy for social interaction since they could not directly observe it. Our data allows us to directly observe and measure social interaction. Moreover, we highlight sentiment contagion as the channel for peer actions to influence investors' decisions. Finally, our paper also adds to the literature by documenting the aggregate impact of peer action on the market.

Most prior works on investor sentiment focus on constructing sentiment indices at the aggregate level to understand asset pricing dynamics. For example, Baker and Wurgler (2013) constructs a sentiment index using market variables such as IPO indicators and trading volume at the aggregate level to explain many asset pricing facts. We expand this literature by providing micro-foundations for sentiment dynamics, with an emphasis on the role of social interaction on propagating the spread of sentiment at the individual investor level.

The remainder of the paper is organized as follows: Section 2 describes the data and explains the main variables used in the analysis. Section 3 documents the sentiment contagion pattern. Section 4 presents the impact of sentiment contagion on individuals' trading decisions and on the aggregate market. Section 5 shed light on bubble episodes from the perspective of sentiment contagion. Section 6 concludes.

2 Data

2.1 Bitcointalk

The primary dataset used in our empirical analysis is from a social network website called Bitcointalk. Bitcointalk was founded in November 2009 for investors to share information about Bitcoin. Compared to other Bitcoin forums, Bitcointalk is special in that its founder Satoshi Nakamoto is the presumed pseudonymous author (or authors) of the original Bitcoin white paper[2] which describes Bitcoin's reference implementation.[3] Endorsed by the founder of Bitcoin, Bitcointalk has become one of the most active online platforms for cryptocurrency investors. As of 3/4/2020, there have been 54 million messages posted for over 1.2 million topics. There are over 2.7 million registered users or investors. For the rest of the paper, we are going to use the two terms "users" and “investors" interchangeably.

In this paper, we focus on conversations on Bitcointalk. Each conversation is presented within one thread, which is a collection of sequential posts where investors can reply to each other and discuss a common topic. Each post is a timestamped message that has a bundle of sentences written by an investor. Most sentences within one post are concise. Figure 1 presents one conversation as an example in our sample. On May 31, 2018, investor DavidLuziz started a conversation by posting the first post of the thread, titled “What does the future of bitcoin look like”, followed by many other investors participating in this thread to share their opinions.

Conversations on Bitcointalk are displayed under different “child boards”, subforums that collect conversations on a common topic. For this paper, we focus on the “speculation” child board, where investors mainly discuss their views on Bitcoin's prospects. The “speculation” child board is an ideal field laboratory for us to study the effect of social interaction on investor sentiment, trading behavior, and aggregate Bitcoin prices and volumes.

We also observe heterogeneity of sophistications among users. Bitcointalk grants users merits, a point system that values those users who contribute to the Bitcoin community.[4] The Bitcointalk website assign

[2]https://bitcoin.org/bitcoin.pdf

[3]Satoshi Nakamoto created the Bitcointalk forum and posted the first message in 2009 under the pseudonym "satoshi"

[4]Specifically, users who publish posts that stand out in quality receive merits. See https://bitcointalk.org/index.php?topic=2818350.0

users with different ranks, from high ranks named "Legendary" to low ranks like "Brand New", depending on their merits and activities. Users with high ranks are typically more sophisticated Bitcoin investors. We therefore categorize users into sophisticated and naive based on these ranks: investors who are legendary users are defined as sophisticated users while the rest users are treated as naive users. In addition, a large fraction of users in our sample choose to voluntarily reveal their age, gender, and their home country. We study heterogeneity in the sentiment contagion effect by these investor characteristics.

2.2 Investor Sentiment

2.2.1 Textual Analysis and Investor Sentiment

In this paper, we use the term "investor sentiment" to refer to investors' beliefs about future Bitcoin prices. We design a methodology to extract investor sentiment for posts collected from the Bitcointalk forum. Since our raw data contains more than 1 million posts, more than we can manually interpret, we employ a standard textual analysis algorithm. Our algorithm is a two-step procedure built on a dictionary of keywords and a natural language processing (NLP) algorithm designed by the Stanford University NLP group.

In the first step, we randomly select 10,000 sentences from our dataset and manually label them into four categories: "positive", "neutral", "negative", and "irrelevant". This follows Baker et al. (2016) and Tetlock (2007). We then use the labeled sentences as our training set and construct a keyword dictionary for each category. For instance, keywords for "positive" category include, but are not limited to, "buy", "increase", and "rise", while keywords for "negative" category include "sell", "decrease", "plunge", etc. The dictionary for the "neutral" category includes "hold", "wait", "unpredictable" and so on. If one sentence does not contain any of the keywords in the above three categories, we label it as "irrelevant". For our analysis, we focus on relevant sentences.

In the second step, we apply the Stanford NLP algorithm to detect sentences that describe the future. One challenge in our labeling exercise is that some posts contain descriptive statements about past performance rather than their beliefs on the future. For instance, a post like "Bitcoin market really increased a lot in the past months" is ambiguous in terms of an investor's opinions about future Bitcoin prices. Relying on the Stanford NLP algorithm, we can identify sentences with forward-looking statements, and label the rest as "irrelevant". As an example, if one sentence contains the keyword "increase", but is associated with a backward-looking statement, then it is labeled as "irrelevant". Our definition of forward-looking statements follows a similar format suggested by the SEC for 10-K reports.[5]

With labels from the above two steps, we can extract investor sentiment from each sentence. We also process sentences with negative particles such as "not" and "couldn't". If a sentence with negative particles has words that fall into the "positive" category, we flip our interpretation and label this sentence as "negative"

[5]Sentences with future tense or terms such as "expects", "anticipates" and so on are categorized as forward-looking statements. See $https://www.sec.gov/Archives/edgar/data/1082027/000139390519000101/neik_10k.htm$ for details.

instead of "positive". We assign a value of 1 to sentences with a “positive” sentiment, -1 for “negative”, and 0 otherwise. For example, “Bitcoin price will roar” is assigned a value of 1. "Bitcoin price is not going to fall" is assigned a value of 1. “Bitcoin is doomed to fall” is assigned a value of -1. “Bitcoin price is really unpredictable” takes a value of 0. To measure the overall sentiment of one post, we take the average of sentiment in all sentences. Therefore, the sentiment measurement for a post has a continuous range from -1 to 1. The out-of-sample accuracy of our algorithm is about 85%.

2.3 Transaction Data

For a subsample of investors, we are able to link their sentiment to their trading behaviors. This group of investors voluntarily published their Bitcoin wallet addresses in a thread officially initiated by the forum organizer. They do so to protect their Bitcointalk accounts from being hacked.[6] Since Bitcoin transactions are all publicly available on the Bitcoin Blockchain, with these published wallet address, we can essentially trace their transactions. We show an example of the Bitcoin trading record in Figure 2. In this figure, we observe both buyer and seller’s Bitcoin wallet addresses, the trade size, and the time of the trade. Of the about 37,262 investors in our sample, we are able to link sentiment to trading decisions for 1284 investors.

2.4 Other Data Sources

To capture common news shocks to the Bitcoin market, we download news about Bitcoin from Ravenpack News Analytics, which tracks news reports about Bitcoin and provides sentiment scores for them from 2011 onward.[7] We refer to the overall tune on Bitcoin in news media as the news sentiment and include it as a control variable in our analysis.

We also collect Bitcoin market information such as returns and trading volume at the hourly level from CoinAPI[8] and calculate daily return volatility. The same data is used in Griffin and Shams (2020) published by CoinDesk.

Moreover, we measure investor attention to Bitcoin with the Google Search Volume Index for keyword “Bitcoin”. Specifically, to capture variation in investor attention to Bitcoin relative to its recent past mean, we define Abnormal Google Search Volume as the difference between the Google Search Volume Index and its past one-month mean, divided by the lagged one-month mean. This procedure follows Da et al. (2011).

[6]This is an official activity organized by the management team at the Bitcointalk forum. When a Bitcointalk account is hacked, its owner can retrieve the account by signing the public wallet address they posted in the thread.

[7]The Entity ID of Bitcoin in Ravenpack News Analytics is A25816.

[8]https://www.coinapi.io/

2.5 Summary Statistics

Our main sample of Bitcointalk conversations starts on May 1st, 2012 and ends on July 30th, 2018. In Panel A, B and C, we provide detailed description for investor sentiment at user, daily and conversation level. In untabulated results, we also document that the average sentiment of a post negatively predicts future Bitcoin returns, up to 30 calendar days. Panel D reports summary statistics for market information of Bitcoin, RavenPack news sentiment and Google Search Volume.

In Panel A of Table 1, we look at how each user publishes posts. For a representative user in our sample, the average sentiment in posts he published is 0.329, which indicates that users in general are optimistic about Bitcoin.[9] However, we do observe a significant fraction of users who are on average pessimistic about Bitcoin. Within the same user, there is significant variation in sentiment over time, with the average with-in user standard deviation of sentiment being 0.604. The distribution for the number of posts is skewed, shown by the gap between the median and the mean number of posts by a user.

Panel B of Table 1 describes post activities within a day. On a representative trading day, the average sentiment is 0.268, which is optimistic and consistent with the statistics at the user level. On a representative trading day, about 129 users participate in conversations and publish around 190 posts.

We can also see from Panel C that posts within one conversation often have dispersed sentiment. The average within-conversation standard deviation of sentiment is 0.648. Users actively participate in conversations. For a representative conversation, 21 users participate with 27 posts. In general, users are responsive in conversations. The median gap between consecutive posts is 0.028 days, about 40 minutes.

Panel D reports summary statistics for variables at the aggregate market level. From the first row of Panel D, we can see that the annualized volatility of Bitcoin is 1727%, considerably higher than the volatility of S&P 500 index. The annualized average return is 160.2%, implying a Sharpe ratio of around 0.926. The second row presents the within-day volatility of Bitcoin using the hourly return data. On average, the volatility within one day is around 4%. We report the volume-related variables for Bitcoin from row 3 to 5. For a typical trading day, there are 11,711 transactions, with 11,158 bitcoins being traded. On average, the daily total trading volume is 23.13 million in dollars. In row 6, we report the summary statistics for the sentiment of RavenPack News, calculated as the average sentiment of news within one day based on the RavenPack database. Of the 2281 days in our sample, we observe news on Bitcoin for 676 days. On average, RavenPack news is optimistic about Bitcoin, with a mean sentiment level of 0.159 and a standard deviation of 0.551. In the last row, we report the summary statistics for the abnormal Google search volume.

We report the correlation matrix between the aggregated sentiment at the daily level and other important variables in Table 2. The correlation between the average sentiment and the number of posts is 0.239, suggesting that users tend to post more actively when they are more optimistic. The average sentiment is

[9]This optimism is consistent with the existing literature on investor expectations (Greenwood and Shleifer (2014), Giglio et al. (2021)), where they also document that investor expectations are on average optimistic.

decreasing in the dispersion of sentiment, increasing in daily Bitcoin return and intra-day Bitcoin volatility. The average sentiment is also positively associated with the RavenPack News sentiment. The post activities align well with investor attention measured by the Abnormal Google Search Volume, with a correlation of 0.184.

3 Social Interaction and Sentiment Contagion

In this section, we investigate whether social interactions propagate the spread of investor sentiment, a pattern we dubbed as *sentiment contagion*. Specifically, we study how an investor' sentiment would be affected by others' sentiment through conversations on the forum.

3.1 Empirical Strategy

We motivate our empirical strategy by an example in Figure 1. In this example, investor DavidLuziz published one post at 05:37:35 AM on May 31st, 2018 (post[0] at the top). After his first post, several other investors joined the thread and had a discussion by sharing their opinions about Bitcoin. At 07:50:10 AM on the same day, DavidLuziz published his second post in the same thread (post[1] at the bottom). As we can see from his two posts, after interacting with other investors, his sentiment changed from neutral to optimistic.

To investigate how sentiment spreads during such social interactions, we study the relationship between an investor's sentiment change from before to after a conversation and the Whisper sentiment sentiment. Formally, sentiment change is defined as the revision in sentiment from post[0] to post[1], while Whisper sentiment is measured as the average sentiment in other investors' posts published between post[0] and post[1] in the same thread. Our hypothesis is that Whisper sentiment positively predicts sentiment change. We also call post[0] and post[1] the prior and the ex-post sentiment, respectively.

To better measure sentiment change in a conversation, we impose that post[0] and post[1] should be consecutive, i.e. the investor does not publish any other posts between post[0] and post[1]. In most of our analysis, we further restrict that the two consecutive posts to be published within a 24-hour window so that sentiment change is less likely affected by other information sources.[10] An investor's consecutive posts could be in two different threads. For these cases, Whisper sentiment is based on posts published between the post time of post[0] and post[1], and are in the same thread as post[0] or post[1]. Our results remain unchanged if we focus on the much smaller subsample with post[0] and post[1] in the same thread.[11]

[10]In Appendix Table A2, we show that our findings are robust to alternative windows such as 12-hour.

[11]See Appendix Table A1

3.2 The Role of Priors

Social interactions could show heterogeneous effects on investors. After all, different users have different priors and therefore respond differently to the opinions in the conversations. Indeed, recent findings in the psychological literature have highlighted a phenomenal pattern called confirmation bias, the tendency to search for, interpret, favor, and recall information in a way that confirms or supports one's prior beliefs or values. Therefore, in our context, we conjecture that users' priors affect how sentiment changes in response to social interactions.

And we find supportive evidence for our conjecture. Specifically, we treat the sentiment level in post[0] as the prior sentiment of the user. By relying on simple OLS regressions, we find that users with different priors indeed respond differently to opinions they receive in the conversation. We show this evidence in Figure 3. Each line in the figure is a fitted line for between sentiment changes and Whisper sentiment conditional on a different level of prior.[12] We have three priors, positive sentiment, negative sentiment and neutral sentiment. As we can tell from the figure, users revise their sentiment upwards more aggressively in response to positive sentiment in Whisper sentiment when their prior is positive, a pattern consistent with the confirmation bias. Formally, we use the seemingly unrelated regression to test the difference in slope for different priors. We find the difference between slopes of positive and negative priors is significant at 1%.

Therefore, to show the impact of social interaction on sentiment contagion in our main analysis, we control for the prior in our main regression setting.[13]

3.3 Sentiment Contagion

Our main specification is

$$SentiChange_{i,j,t_0\to t_1} = \beta_1 WhisperSentiment_{i,t_0\to t_1} + PriorSenti_{i,t_0} + \gamma Control_{t_0\to t_1} + \alpha_i + \gamma_{d(t_0)} + u_{i,t_1} \tag{1}$$

where $SentiChange_{i,j,t_0\to t_1}$ is the sentiment change from Post[0] to Post[1] by investor i in the conversation j . Timestamps for Post[0] and Post[1] are denoted as t_0 and t_1, respectively. $WhisperSentiment_{i,t_0\to t_1}$ is the Whisper sentiment that investor i has received between her consecutive posts from time t_0 to t_1. To address the omitted variable issue due to the confirmation bias, we also control for $PriorSenti_{i,t_0}$ which is the prior sentiment of investor i at time t_0. We include both user fixed effects and date fixed effects, denoted as α_i and $\gamma_{d(t_0)}$. For $Control_{t_0\to t_1}$, we include control variables for news shocks, market fluctuations and the aggregate sentiment of posts at the Bitcointalk community level. We cluster standard errors at user and date levels.

If social interaction propagates sentiment contagion, then after participating in a conversation, investors

[12] For the purpose of comparison on their slopes, we ignore the intercept.

[13] If we regress sentiment change on Whisper sentiment without controlling for prior sentiment, we face an omitted variable bias.

should revise their sentiment towards the direction of the overall sentiment in the conversation. Therefore, we should anticipate the coefficient β_1 of Whisper sentiment to be statistically significant and positive. The result in Table (3) confirms our conjecture. Column (1) reports the result from regressing sentiment change on Whisper sentiment after controlling for the priors. As the average level of Whisper sentiment increases, investors update their sentiment more aggressively towards the sentiment in the conversation: all else being equal, one standard deviation increase in Whisper sentiment is associated with an increase in changes in sentiment of 1.332%.

Sentiment change from post[0] and post[1] can also be driven by common news shocks that simultaneously arrive during the conversation. This story, with a similar spirit in Feng and Seasholes (2004) based on Chinese data, may contaminate our findings. To rule out this possibility, we include several controls in column (2) of Table 3. We first include the contemporaneous news sentiment variable from Ravenpack News Analytics to directly proxy for the arrival of information within the consecutive pairs. Specifically, we calculate the average of sentiment in Ravenpack news as controls for news arrivals between two consecutive posts. We find its coefficient is positive but insignificant.

An alternative explanation related to common news shocks is that investors may respond to news with a delay, and there is a further delay before they post messages. In this situation, some news arrives before post[0], but the primary investor only gets access to it after post[0] was published (but before post[1] was published). Then sentiment change and Whisper sentiment may be driven by outdated news that arrived before post[0]. To address these concerns, we add lagged Ravenpack news shocks (up to 48 hours before post[0]) as additional controls. The effect of Whisper sentiment remains robust, indicating that our findings are not driven by a channel of lagged responses to news shocks.

News sentiment in Ravenpack may not fully capture the arrivals of contemporaneous news shocks. In column (3), we further control for other types of contemporaneous information sources, mainly the market variables such as Bitcoin return, volatility and total number of trades over the time interval of the conversation. The semi-strong Efficient Market Hypothesis implies that these market variables should reflect all possible contemporaneous news shocks. Therefore, by including these controls, we set a higher bar to study the impact of social interactions on sentiment change. We find that the regression coefficient of Whisper sentiment remains unaffected. In the Appendix table A3, we further control for news arrivals in the past 7 days and our findings remain unchanged. Hence, the impact of Whisper sentiment on sentiment change remains robust after accounting for a story of common news shocks. Notice that our findings do not deny the effect of the market information — we find the coefficient for Bitcoin return remains statistically significant.

Thus far, our analysis focuses on sentiment contagion through direct social interactions –sentiment spread across investors when they participate in the same conversation. It is possible that sentiment may also spread via indirect social interactions, wherein investors may simply browse other posts in Bitcointalk while not directly participating in the conversation by publishing a post. Although it is not our main focus, we still present preliminary analysis to test these indirect social interactions. Specifically, we define a variable called forum sentiment, which captures the average sentiment of posts on the Bitcointalk forum published between timestamps of post[0] and post[1], but not in the same threads of post[0] and post[1]. The forum sentiment variable measures the overall sentiment outside of the conversations the user participates in. We include it in our regression as an additional control. Column (4) reports a negative coefficient for forum sentiment, which indicates that this indirect social interaction may not be the main channel for sentiment contagion.

3.4 Discussions on Other Omitted Variable Issues and Reverse Causality

By including date fixed and investor fixed effects, we address several omitted variable issues. For instance, for the observed variations in news that are not captured by news sentiment in Ravenpack or other news sources mentioned above, our date dummy would absorb the majority of fluctuations at the aggregate level. Moreover, adding date-fixed effects also helps us address important identification issue related to reverse causality. As an example, if changes in investor sentiment are driven by some unobserved time-varying component such as the aggregate sentiment level in the Bitcoin market, Whisper sentiment and sentiment change could be mechanically correlated: sentiment changes increase when the aggregate sentiment in the market increases, which is also associated with an increase in Whisper sentiment. Including date-fixed effects helps us rule out this possibility, because the effect of the Whisper sentiment in our setting is estimated from cross-sectional differences in sentiment change and variation of those cross-sectional differences over time, not from the aggregate time-variation.

Controlling for investor fixed effects helps us eliminate alternative stories explained by time-invariant characteristics at the user level, such as gender and IQ, factors that are shown to influence individuals' decision making in previous studies. [14] Moreover, since we restrict a 24-hour window on consecutive pairs, most time-varying characteristics such as education that change at an infrequent level are unlikely to explain the sentiment contagion pattern. Our results remain unchanged if we choose alternative lengths for this short window.

3.5 Placebo Test

To address concerns on spurious trends in investors' sentiment change over the course of conversations, we provide a placebo test. Specifically, this placebo test shows how sentiment changes are affected by the average sentiment in a random conversation that happens between the timestamps of post[0] and post[1] but the investor does not participate. If the sentiment contagion pattern we have documented is valid, then the average sentiment in a contemporaneous but random conversation should not influence investors' change in

[14] For example, see Barber and Odean (2001), D'Acunto et al. (2019), Grinblatt et al. (2011)

sentiment. We find supportive evidence from this placebo test.

The results are presented in Table (4).Increases in sentiment of random conversations do not predict an upward revision in sentiment. Therefore, the placebo exercise confirms that the regression tests of social interaction on sentiment contagion are unlikely to give false-positive results. Moreover, it highlights the important role of social interactions in sentiment contagion: without social interactions, sentiment contagion disappears.

3.6 Heterogeneity by User Features

The impact of social interactions on sentiment contagion depends on user features. We demonstrate such heterogeneous effect across four dimensions: user sophistication, gender, language skills and age.

Classical theories on investor sentiment typically assume there are two types of investors: naïve investors who are sentimental and are often driven by psychological factors and sophisticated investors who serve as the counteracting force and arbitrage away the mispricing caused by investor sentiment (De Long et al. (1990), Lee et al. (1991), Barber and Odean (2013)). On Bitcointalk, investors are classified into legendary and non-legendary categories based on their experience and contribution to the Bitcointalk community. Therefore, we study the heterogeneity in sentiment contagion for these two types of investors. Columns (1) of Table (5) presents the heterogeneous effect related to investor sophistication. Specifically, we create a naive investor dummy and interact it with Whisper sentiment. We find that this interaction term is positive and statistically significant, confirming that naive investors tend to be more affected by sentiment in conversations. Compared to a sophisticated investor, a naive investor is 4.1% more likely to be affected by sentiment in conversations. Our findings match the standard dichotomy in the literature: increased investor experience and sophistication indeed make investors less susceptible to sentimental elements in the market.

Previous findings have also highlighted the role of gender on investor behaviors. [15] Motivated by these pioneering works, we study the heterogeneous impact of social interaction related to gender. In column (2) of Table 5, we use the interaction term between the female dummy and Whisper sentiment to predict sentiment changes. The coefficient is negative and significant under the most stringent regression configuration in our setting. Compared to male users, female users are 9.3% less affected by sentiment in conversations, but the gap is statistically weak. If we interpret sentiment contagion partially as a result of overconfidence in the sentiment in conversations, our findings are in line with the previous findings in the literature.

In column (3) of Table 5, we study whether a native English speaker is affected less by sentiment in conversations. The speculation subforum on Bitcointalk is discussed in English. Therefore, we hypothesize that users whose mother tongue is English may process the information in threads more efficiently, therefore are less misguided by the sentiment in conversations. We find it is indeed the case. The interaction

[15]Psychological literature documents that men are in general more overconfident than women. (Lundeberg et al. (1994)). Following the spirit in this paper, Barber and Odean (2001) document that men trade 45 percent more than women.

term between the native-English-speaker dummy and Whisper sentiment is negative and significant. The magnitude of coefficient is 0.155, indicating that native English users are 15.5% less likely to be affected by sentiment in conversation.

In column (4), we investigate the role of age in sentiment contagion. We create a age dummy for users older than 40 years old, and find no effect. In untabulated results, we use alternative age dummies with different cutoffs and our results remain unchanged. Our findings indicate that sentiment contagion does not seem to relate to age.

3.7 Heterogeneity by Conversation Features

Conversations with distinct features may influence sentiment contagion in different ways. Our empirical findings confirm this conjecture. Specifically, we focus on three features of conversations: whether a conversation contains at least one sophisticated investors, how many posts are published in a conversation and how disperse attitudes are in a conversation.

Table 6 presents our results. In column (1), we create a dummy for conversations evolved by sophisticated investors and interact it with the Whisper sentiment. We find that conversations with at least one sophisticated investor are more persuasive in influencing users' sentiment revisions. All else being equal, such conversations are 5.5% more likely to affect users' sentiment revision. We interpret this finding as meaning that users put more trust in conversations pushed forward by sophisticated users. In the literature, researchers typically treat sophisticated investors as those who make wise decisions, but usually remain silent on the externality of their actions. Our findings here indicate that sophisticated investors' behaviors affect other investors' decisions.

In column (2), we examine how users are influenced by the number of posts in a conversation. Specifically, we interact the total number of posts with the average sentiment in the conversation to predict sentiment changes. We find that the interaction term is statistically significant and positive, an indication that conversations with more posts lead to stronger sentiment contagion.

There could be at least two ways of interpreting our findings in column (2). One explanation based on Bayesian learning is that as investors observe more information from an increased number of posts, they face less of uncertainty and therefore become more affirmative in revising their sentiment. Alternatively, another way to interpret our finding is based on one prevailing pattern in psychology literature: "illusory truth effect". This pattern indicates that individuals tend to believe false information to be correct after repeated exposure to that information.[16] Although the irrational feature of sentiment indicates that our findings may be more susceptible to the latter, we do not aim to disentangle these two explanations and leave it for future research. Either way of interpreting our findings in column (2), the pattern is quite robust and sheds light on our later

[16]One explanation for such psychological pattern is that repetitively hearing that a certain fact increases familiarity and that familiarity can overpower rationality when forming beliefs. See Hasher et al. (1977).

discussions of bubbles.

We also study how attitude dispersion in a conversation affects sentiment contagion. Conditional on the same level of Whisper sentiment, with increases in attitude dispersion, there could be two counteracting forces on sentiment contagion. On the one hand, as users see a richer set of information, they might be more resolute in revising their sentiment. However, dispersed attitudes may simultaneous reduce user's confidence in the accuracy of the information in posts, and therefore become more hesitant to change their sentiment based on this conversation.

To investigate which force dominates, we construct a measurement for attitude dispersion in column (3). Specifically, we calculate the attitude dispersion as the standard deviation of sentiment of posts published between post[0] and post[1] in a conversation. In order to get a meaningful standard deviation, we require there to be at least two posts between post[0] and post[1]. Conditional on the same level of Whisper sentiment, conversations with more dispersed attitude have less impact on sentiment contagion. More dispersed attitude may lead users to become less confident on accuracy of the information, and therefore would be less willing to revise their sentiment after participating the conversation.

Changes in the number of posts happen simultaneously with variations in attitude dispersion. To disentangle their independent impact on sentiment contagion, we include two interaction terms with Whisper sentiment, one for the number of posts and the other for attitude dispersion in column (4). We find the pattern is quite robust as we show in column (2) and (3) and each channel independently exerts impact on sentiment contagion. All else being equal, adding one more post makes the conversation 0.4% more likely to affect users' sentiment. Moreover, a one standard deviation increase in attitude dispersion makes the conversation $0.162 \times 0.172 = 2.786\%$ less likely to influence revisions in sentiment.

4 Real Effects of Sentiment Contagion

If social interaction leads to sentiment contagion, how would such sentiment contagion affect investor trading behaviors and influence the Bitcoin market? In other words, should we anticipate any real effects of social interaction? On the one hand, changes in sentiment may predict investors' future trading behaviors, as indicated by some recent studies that there is a significant link between investors' attitudes and their transactions (Giglio et al. (2020)). However, changes in sentiment may not have any real effects at all, because investor sentiment is transient and closely related to animal spirits, and such irrationality can be corrected away by arbitragers (Baker and Wurgler (2013)). In this section, we examine the real effects at both individual level and aggregate market level. [17]

[17] A growing literature has documented that investors' financial decisions respond to the experience or information set of their friends. However, the underlying mechanism remains an open question.Kuchler and Zafar (2019) document that the impact of friends' experience on one's real estate investment decisions may be explained by changes in beliefs. Heimer (2016) document that investors' preferences may be influenced by their social network and thus they may display a stronger disposition effect. We aim to highlight

4.1 Impact at the Individual Investor Level

To evaluate the real effect of sentiment contagion at the individual level, We first link changes in sentiment to users' trading decisions. Following our main empirical setting in Table 3, we continue to focus on users' consecutive posts and link such sentiment changes to their future trading activities. We conjecture that social interaction will induce investors to not only change sentiment, but also to trade. To formally examine it, we run the following regression:

$$Buy_{i,t+k} = \beta_0 + \beta_1 SentiChange_{i,t} + FE + \Sigma_m \gamma_m Control_{i,t,m} + u_{t+k}. \tag{2}$$

The independent variable, $SentiChange_{i,t}$,is defined as the sentiment change between the consecutive posts post[0] and post[1] within 24 hours.[18] To exclude other alternative drivers for individuals' trading behaviors, we include proxies for news shocks, market fluctuations and the forum-level aggregate sentiment as additional controls. In all specifications, we include both user and date fixed effects to subsume unobserved user- and date- level variations. We cluster standard errors at user and date levels.

For the dependent variable, we use the dummy $Buy_{i,t+k}$ that equals one if user i's first transaction after post[1] is a buying decision and such transaction occurs within a specific window of length k after the timestamp t of post[1]. We focus on the *first* transaction rather than a sequence of transactions after post[1], because we believe that it is the most direct measurement in our setting to capture the impact of sentiment change on trading. Should we use alternative measurements such as the cumulative trading volume of subsequent transactions after post[1], we face a problem that these subsequent transactions may be affected by users' future sentiment changes driven by information arrivals later on. If it were the case, the relation between sentiment change and the resulting trading decisions would be contaminated. For the length of the gap between post[1] and the first transaction,on the one hand, we need to give the user enough time to adjust the position, after all, it takes a certain amount of time to place an order; but on the other hand, this window should not be too long, because a too long window may cause investor sentiment to be affected by other confounding factors such as market fluctuations during this period. Therefore, we choose a 12-hour window for our analysis in the main setting. We also use an alternative window with infinite length to show how our results are affected by the choice of the window.

The first two columns in Table 7 presents our main findings with the 12-hour window. Overall, as investors change their sentiment after social interactions, they are also likely to make a transaction that is consistent in direction with that of their sentiment changes. In column (1), without any additional control variables, a one unit level increase in sentiment makes investors 0.202% more likely to buy Bitcoins in the subsequent 12-hour window after post[1]. Such a predictive pattern holds even after controlling for news shocks, market fluctuations and forum sentiment, although the statistical significance weakens slightly as

the sentiment contagion channel whereby social interactions affect investment decisions by changing investor sentiment.

[18]Appendix Table A5 report the results using an alternative window of 12 hours and the findings remain unchanged.

shown in column (2).

In column (3) and (4), we expand the length of window from 12 hours to infinite, and study how sentiment changes unconditionally predict the direction of the first transaction after post[1]. As shown in column (3), the statistical power of sentiment change drops remarkably but the sign of the coefficient remains positive. This finding is not surprising: if the first transaction happens in the distant future, then it's very likely driven by users' future sentiment changes that are contaminated by future information arrivals and therefore may not be correlated with sentiment changes between post[0] and post[1]. Therefore, it is difficult to capture the long-term real effect of sentiment changes between post[0] and post[1]. Another point worth noting is that the magnitude of the coefficient for sentiment change increases, indicating that the overall probability of buying after post[1] is more likely after seeing an increase in sentiment. This happens because we now consider a longer window, and more buying decisions would appear as the first transaction in this longer window.

Overall, our results imply that changes in investor sentiment predict the occurrence and the direction of individuals' trading in the short run. Sentiment changes induced by social interactions are not sideshows. They significantly influence investors' trading decisions.

The role of belief changes on trading decisions is also documented in recent studies such as Giglio et al. (2020). They find that, since trading occurs infrequently, belief changes cannot predict the occurrence of trading (the extensive margin), but can explain the direction of trading once a transaction is made (the intensive margin). In comparison, we document that sentiment changes do predict trading both on the extensive margin (occur within 12 hours) and on the intensive margin (direction of the trading). The main difference is that we document a stronger link between sentiment change and trading on the extensive margin. A detailed investigation into such discrepancy is beyond the scope of the paper, here we provide a brief discussion on the potential causes of such difference. First, in terms of the frequency of observations, they rely on a bimonthly survey to extract beliefs and link to trading records. We rely on posts whose timestamps can be down to the millisecond to extract sentiment. Our rich data help us fully capture the links between sentiment change and trading decisions that may be missing in a data set whose time dimension is of lower frequency. Second, in terms of investor profile, they study wealthy investors who seemingly make careful decisions. We study investors who are very sentimental and more subject to behavior biases such as overconfidence. If investors are overconfident about their attitude, they may be more likely to make a trading decisions based on their attitude.

4.2 Impact at the Aggregate Market Level

Building on the fact that investor sentiment induces individuals to trade, we go one step further and analyze the aggregate impact of sentiment changes triggered by social interactions on market dynamics. Specifically, using the unique dataset from Bitcointalk, we construct a daily indicator for the intensity of sentiment contagion at the aggregate market level. [19] Different from traditional indicators which usually calculate the average level of investor sentiment, the constructed Sentiment Contagion Indicator (or SCI) captures the spread of sentiment and therefore captures investor sentiment in a dynamic sense. Another unique feature of SCI is that it is a sentiment measurement that is highly associated with social interactive activities. Existing literature on social interaction focus on its impact on individual investors' decision making, and little work has been done, according to our knowledge, to analyze its impact at the aggregate market level.

4.2.1 Construction of SCI

Our objective is to capture the daily intensity of sentiment contagion that happens when investors directly interact with each other. The general strategy we follow has two steps. First, we focus on the affected investors whose sentiment changes after participating in a conversation within a given day. To make sure that their sentiment changes are induced by conversations, we also restrict the direction of sentiment change to be of the same sign with the overall sentiment in the conversation. Therefore, the number of affected investors measures the intensity of sentiment contagion: as the number of affected investors increases in the market, the intensity of sentiment contagion triggered by social interactions also increases.

One caveat is that the number of affected investors can also be influenced by the scale of the investor pool that uses the Bitcointalk website. As Bitcointalk builds up its reputation overtime, the total number of active investors in the forum may also increase over time. This is indeed what we find. Therefore, in our second step, we eliminate the time trend and seasonality by regressing the series for the number of affected investors on the weekday dummy and the year-month-pair dummy. We use the residual as our SCI index. The SCI index has a standard deviation of 13.125 and a mean of 0.

4.2.2 SCI and Future Trading Volume

The key finding in section 4.1 is that sentiment contagion induced by social interactions affects individuals' trading decisions, especially for the window within 12 hours after the sentiment change. Since SCI captures the aggregated intensity of sentiment contagion, a natural conjecture is that the SCI index may predict and explain future trading volume in the aggregate market level. To test our conjecture, we link the SCI index to the future trading volume by running the following regression:

$$abvolume_{t+1} = \beta_0 + \beta_1 SCI_{i,t} + \Sigma_m \gamma_m Controls_{t,m} + u_{t+k} \tag{3}$$

[19] One assumption behind our approach is that the dataset we use (Bitcointalk) is good enough to represent the overall trend in the Bitcoin market. In no way we are claiming that this Bitcointalk forum contains the majority of Bitcoin investors. However, we do assume that the Bitcointalk forum captures very well the *trend* of the social interactive activities over time. Considering the influence of Bitcointalk website among Bitcoin investors and the fact it was founded by Satoshi Nakamoto, who is the founder of Bitcoin, we believe that Bitcointalk does reflect the trend in the market.

where $abvolume_{t+1}$ denotes the abnormal trading volume of Bitcoin on day $t+1$. Specifically, to handle its persistence, we normalize the raw trading volume by subtracting the mean level of trading volume in the past 14 days. The key explanatory variable is $SCI_{i,t}$. For control variables $Controls_{t,m}$, we calculate the average levels of sentiment in Ravenpack news in the past 14 days to proxy for news arrivals. We also control for market information by adding two Bitcoin market variables: Bitcoin Volatility and Number of Transactions, each of which is the cumulative sum in the past 14 days. We control for forum sentiment by calculating the average sentiment of posts on the Bitcointalk forum published in the past 14 days.

Column (1) and (2) of Table 8 present our findings. Overall, an increase in the intensity of sentiment contagion significantly predicts the abnormal trading volume in the next day. For the results based on the univariate regression in column (1), a one percentage point increase in SCI increases the abnormal trading volume in the next day by 22.804 basis point (in millions). This result remains strong even after controlling for news shocks, market fluctuations and forum sentiment in the past 14 days.

4.2.3 SCI and Future Volatility

A long strand of literature starting from Black (1986) has been interested in the relation between investor sentiment and the volatility of asset prices. If investors base their trading decisions on sentiment, then changes in sentiment lead to more noise trading and excessive volatility. Recent studies (Da et al. (2014), Antweiler and Frank (2004)) find empirical support for this relation by using the U.S. stock market data. In this section, we study the role of social interactions in linking investor sentiment and market volatility.

As we have shown in previous sections, changes in sentiment of Bitcoin investors in our data set lead to more noise trading. We therefore conjecture that the intensity of sentiment changes triggered by sentiment contagion, measured by SCI, should positively predict future volatility in the market. To test this conjecture, we run the following regression:

$$\ln(rv_{t+1}) = \beta_0 + \beta_1 SCI_t + \Sigma_m \gamma_m Control_{t,m} + u_{t+k} \tag{4}$$

We calculate the daily realized volatility as the standard deviation of hourly returns and normalize it to get the annualized volatility rv_{t+1} for date $t+1$. The key explanatory variable of interest and the control variables are identical to those in equation (3).

Columns (3) and (4) of Table8 present our results. The future volatility of Bitcoin returns in the next day loads positively and significantly on the SCI indicator. For example, in column (3), a percentage point increase in SCI corresponds to a 2.734 increase in the realized volatility of Bitcoin returns in the following day. Including additional controls in does not drive away our results, as indicated in column (4). This is consistent with our conjecture: as social interactions trigger more sentiment changes, more investors execute noise trading which pushes up the volatility of returns in the market. We also find that return volatility loads positively on its realizations in the past 14 days, which indicates a high level of persistence.

4.2.4 SCI and Future Returns

A small number of papers in the literature have examined the predictive power of online message boards on future asset returns. For example, Antweiler and Frank (2004) document that higher message postings predict subsequent stock returns, but their results are economically small.

Consistent with Antweiler and Frank (2004), we do not find significant predictive power of the SCI on future returns (shown in column (5) and (6) of Table 8). One explanation is that SCI represents the sum of changes in demand for Bitcoin in both ways: some investors are affected by sentiment contagion to buy while others to sell. At the aggregate level, SCI does not predict which way dominates the other and therefore does not predict future price movement.

5 Sentiment Contagion and Bubbles

Bubbles have long been recognized as episodes featuring elevated investor sentiment and iirational exuberance, along with rapid rise in both asset pricing and trading volume. Researchers are fascinated by not only these features alone, but also the underlying mechanisms behind them. The Bitcoin market provides a valuable opportunity to investigate bubbles. Indeed, the Bitcoin market has been often associated with bubbles since its inception. In the past decades, the Bitcoin price has experienced several drastic changes in price that are not easily explained by its fundamentals, which inevitably makes people pay attention and be curious about its bubble features. Moreover, Bitcoin was born in an age with various developed social platforms, and the process of its popularity is also closely related to the social network. Therefore, in this section, we focus on bubble-like episodes in the Bitcoin market. Our data allows us to provide evidence on the time-varying feature of social interactions and sentiment contagion during bubble episodes. We aim to gain a more accurate understanding of some features of bubbles from the perspective of social interaction and sentiment contagion.

5.1 Identify Bitcoin Bubbles

We define each bubble by identifying the horizon of the days when the Bitcoin price reaches its local peaks and then bursts afterwards. Specifically, we search for the days when there are at least 100% increase of Bitcoin returns during the past one month and 40% decline in Bitcoin value within the subsequent one month. It suggests that we could identify consecutive days. For example, if we identify a day of March 1, 1999 as a day around its local peaks, the day of March 2, 1999 may also qualify. In this case, we merge them together to get a longer window. We pick the day with highest Bitcoin price from this merged window of days as the peak day of each bubble. The bubble formation episode is then defined as a continuous window of days before its peak day, and we require that the past 30-day return of each day in this window be higher than 30%. The bubble burst episode is defined as the one-month window right after the peak day (with a return of less than 40%).[20]

Our choice of 100% return is meant to conform to Fama and others' notion that a bubble, if it exists, begins with a large price run-up. A return threshold of 100% is able to pick up most episodes that anecdotal evidence has suggested were Bitcoin bubbles, such as the episode around the end of 2017. It is also worth noting that we define the bubble from an ex-post perspective by relying on future information. This approach does not harm our analysis since we do not aim to predict bubbles. The cutoffs we rely on, the 100% increase and 40% crash, follow a similar spirit in Greenwood, Shleifer and You (2018). The identified set of bubbles are very robust to these cutoffs. We identify two bubble episodes within our sample horizon. The first one starts from January 10, 2013 to May 9, 2013, with the peak day on April 9, 2013. The second one starts from September 20, 2017 to January 17, 2018, with the peak day on December 18, 2017.

In Table 9, we report the key features of the identified bubble episodes and compare it to the non-bubble episodes in the Bitcoin market. In Panel A, we report the summary statistics for market variables. We find that, compared to the non-bubble episodes, the daily Bitcoin return (Annualized) is more than 8 times higher in the identified bubble episodes and the return volatility within a day also increases by 55.556%. The total number of transactions almost doubles and the total dollar volume grows by 2.5 times. If we look at the news report, the RavenPack news sentiment jumps from 0.149 to 0.301, more than double of the level in the non-bubble episodes. Abnormal Google Search volume surges from below zero to 0.324. In summary, consistent with the definition for bubble, our identified episodes feature rapid rises in returns, volume and market fluctuations, along with increasingly pervasive optimism in the media coverage and explosive investor attention.

Sharp changes in market conditions naturally affect the content and attitudes in people's conversations about Bitcoin. In Panel B, we summarize the key variables on investor sentiment and the social interactive activities. We find that during bubble episodes, users are more optimistic — investor sentiment increases by 0.027/0.265 = 10.189%— and disagreement among users also declines—the standard deviation of sentiment drops by 0.01. Such decrease in disagreement is remarkable, especially given the dramatic 76.300% increase in the total number of users participating in the conversation during the bubble. We also witness a great wave of new and naive investors joining the market, consistent with descriptions for bubble episodes in Shiller (2015). Not surprisingly, the total number of posts within a day also rises sharply by 68.2%. Moreover, the number of posts with positive sentiment increases disproportionately: the fraction of posts with positive sentiment increases by 0.024. The difference between these features in Panel A and B are statistically significant. We also rely on the seemingly unrelated regression to jointly test the significance of these feature differences. The joint F-statistic and its corresponding p-value support that bubble and non-bubble episodes

[20]For example, one bubble we identify has the peak day of December 18, 2017. Moving backward from the peak day, the first date with a past 30-day return being less than 30% is September 19, 2017. Then we identify the bubble formation episode as the days between September 20,2017 and December 18, 2017. The burst episode is from December 19, 2017 to January 17, 2018.

are significantly different, which help justify our approach of identifying bubbles.

5.2 Self-enhancing Transmission Bias

How do factors in the market drive the changes in investors' conversations? In this section, we answer this question by investigating individual investors' post decisions, or the "sending schedules" dubbed by Han et al. (2022). Since the focus of most conversations on Bitcointalk is about Bitcoin prices and returns, we investigate how recent returns influence their post decisions. In spirit, our analysis is consistent with the implications from Han et al. (2022) as they also model investors' decision to discuss their strategy as an increasing function of recent returns.

Table 10 presents our findings. For all market variables in this table, we use the 14 days as the rolling window to calculate the cumulative sum. Our results remain very robust using alternative choices of length for the window (for example, see Appendix Table 10). For presentation purposes, the reported coefficients are all multiplied by 100.

In the first two columns, we study how recent recent Bitcoin returns determine investors' decision to make a post. We focus on each individual's active horizon between the day they register on the Bitcointalk forum and the end date of our sample. We construct a dummy variable for each investor-day pair, which equals one if the investor publishes at least one post on that day and zero otherwise. In column (1), we do not include any control variables except for their priors. The coefficient of returns is 0.166 (t-value = 3.96), suggesting that a one percentage point increase in recent returns increases the likelihood of publishing a post by 0.166 percentage point. The unconditional probability of writing a post is 0.446 percentage point. A 0.166 percentage point increase is economically significant. In column (2), we include controls for recent news sentiment, market variables and forum sentiment. The coefficient of recent return remains unchanged and statistically significant.

In column (3) to (6), we investigate, conditional on making a post, how recent returns affect the post sentiment. In column (3) and (4), the dependent variable is a dummy for posts with positive sentiment and zero otherwise. We do not include control variables in column (3). The coefficient of recent returns is 5.807 (t-value = 9.25). It indicates that a one percentage point increase in recent returns pushes up the probability of posting optimistic opinions by 5.807 percentage point. The unconditional probability of a post with positive sentiment is 52.962 percentage point. A 5.807 percentage point increase is economically sizable. In column (4), we add controls for recent news sentiment, market variables and forum sentiment. The coefficient of return still remains statistically significant (t-value = 3.90). In column (5) and (6), we construct a dummy for posts with pessimistic sentiment, and find that a one percentage point increase in recent returns reduces the probability of posting pessimistic sentiment by 4.274 percentage point. Together, our findings indicate that increases in recent returns induce investors to publish more posts with positive posts.

Results in Table 10 confirm the self-enhancing transmission bias proposed in Han et al. (2022): when

Bitcoin market performs well recently, investors are more likely to write posts and they tend to disproportionately publish more posts with positive sentiment.

Such self-enhancing social transmission bias explains the pervasive optimism in bubbles documented in Table 9, from two perspectives: a rapid increase in the total number of posts and a significant increase in the fraction of posts with positive sentiment which leads to a reduction in attitude dispersion across users. The increased fraction of posts with positive sentiment is further confirmed in the appendix Table A4. We multiply each coefficients by 100. As shown in column (1), when recent Bitcoin returns increases by 100%, the fraction of positive posts within one day increases by 2.978% (t-value= 3.53). In contrast, the fraction of posts with negative sentiment on Bitcoin decreases significantly by 2.059% (t-value = 4.00). The impact of recent Bitcoin returns is not driven away if we include control variables.

5.3 Sentiment Contagion in Bubbles

Each of the two features of the pervasive optimism indicates an elevated intensity of sentiment contagion during bubble episodes. First, increases in the number of posts in a conversation make sentiment contagion more likely to occur, a pattern we document in column (2) of Table 6. Similarly, following the pattern in column (3) of Table 6, attitude dispersion in conversation also drops sharply in bubbles, which further pushes up the intensity of sentiment contagion. Moreover, aside from the two channels that directly result from self-enhancing transmission bias, the increase in the fraction of new investors is also going to foster the spread of sentiment. This happens because new investors are more easily affected by conversations as we show in Table 5.[21]

Altogether, we anticipate an increased intensity of sentiment contagion during the bubble episodes, and we confirm our hypothesis in Table 11. Specifically, we create a bubble dummy for the bubble formation episodes and interact it with the Whisper sentiment. The coefficient of this interaction term then captures the elevated intensity of sentiment contagion during the bubble episodes. The findings in column (1) confirm our conjecture: during bubble episodes, the intensity of sentiment contagion increases by 3.00 percentage point, which is economically sizable to the intensity of 5.20 percentage point in the non-bubble episodes. If we alternatively compare the bubble episodes to non-bubble and bubble burst episodes, we get very similar results, indicating that our findings is not affect by the choice of the benchmark episodes. Such increase in sentiment contagion results from a combination of salient features in bubbles—a rapid increase in the total number of posts, a reduction in attitude dispersion across users and a rising number of new users—and many of these features are associated with the self-enhancing social transmission bias.

[21] Increasing in the number of new investors may simultaneously make the conversation less creditable, as documented in the column (1) of Table 6. However, in untabulated analysis, we show that the fraction of conversations with at least one sophisticated user does NOT drop in bubbles. Therefore, increasing the number of new investors does not trigger the effect documented in column (1) of Table 6 to reduce sentiment contagion.

5.4 Sentiment Contagion and Trading Volume in Bubbles

Why do investors trade so much in bubbles? Explaining the high trading volume has been an intriguing yet challenging task for researchers.[22] In this paper, we provide a new perspective on understanding the high trading volume in the bubble formation episode: the elevated intensity of sentiment contagion. The intuition is a direct extension from our previous findings. We have documented that sentiment contagion in social interactions leads investors to trade and further predict the trading volume at the aggregate market level. Since sentiment contagion become more intensive during bubble episodes, as a result, we should anticipate a high trading volume in bubble episodes.

This is indeed what we find in the data. In Figure 4, we plot the time trend for both trading volume and the SCI indicator in two identified bubble episodes. For trading volume, we use the abnormal trading volume constructed from the raw trading volume by subtracting its mean level in the past 14 days. For the SCI indicator, we use the one constructed in the previous section 4.2. We see a very synchronous pattern between the two variables. In the upper panel we plot the bubble formation episode from January 2013 to May 2013. When the SCI indicator increases, the trading volume the next day tends to increase sharply. The correlation coefficient is statistically significant at 0.627(p-value = 0.00). In the lower panel, we plot the bubble formation episode from September 2017 to January 2018. The correlation coefficient is 0.259 (t-value = 0.00).

6 Conclusion

Echoing calls to "move from behavioral finance to social finance" (Hirshleifer (2020)) and towards a better understanding of "the epidemiology of narratives" (Shiller (2017)), this paper provides direct evidence for the role of social interaction on sentiment contagion. By relying on textual analysis techniques to extract investor sentiment on an influential online investment community, we document a strong sentiment contagion channel whereby investor sentiment on Bitcoin spreads through conversations. Moreover, sentiment changes induced by social interactions influence individual investors' trading decisions and predict future market volatility and trading volume.

Furthermore, we provide new insights on bubbles from a perspective of sentiment contagion. Following Fama's notion about the bubble, we identify episodes with salient bubble features by focusing on price run-ups and crashes. We then rely on the self-enhancing transmission bias to explain the pervasive optimism during bubble episodes — we observe that investors participate more actively in conversations, investors are more optimistic and the dispersion in attitudes among investors decreases. This pervasive optimism leads to an elevated intensity of sentiment contagion and explains the high trading volume in bubbles.

We believe that this paper makes a strong case for the role of social finance in understanding investors'

[22] See Barberis et al. (2018), DeFusco et al. (2017), Liao et al. (2021).

decision-making and market dynamics. The documented sentiment contagion through conversations not only sheds new insights on sentiment, but also highlights the potential of social interactions in explaining an array of finance phenomena. Moreover, the uniqueness of our data allows us to provide direct evidence on several important but inconclusive patterns, such as the self-enhancing transmission bias. Finally, the data set of the online investment community can be an ideal platform to study other important questions in the field of social finance, and we leave these investigations to future research.

Figure 1: Empirical Strategy: An Example

This figure presents an example for our empirical strategy. We show how we measure sentiment changes by calculating the difference between sentiment in post[0] and post[1], and how we measure Whisper sentiment by calculating the average sentiment of others' post in conversations. Each conversation is referred to as a thread with a topic. In the conversation, users can interact with others by posting.

Figure 2

This figure presents an example of an Bitcoin transaction record on the Blockchain. For each transaction, we are able to observe the wallet address for both the sellers and buyers, the corresponding transacted amounts and the timestamp of the transaction.

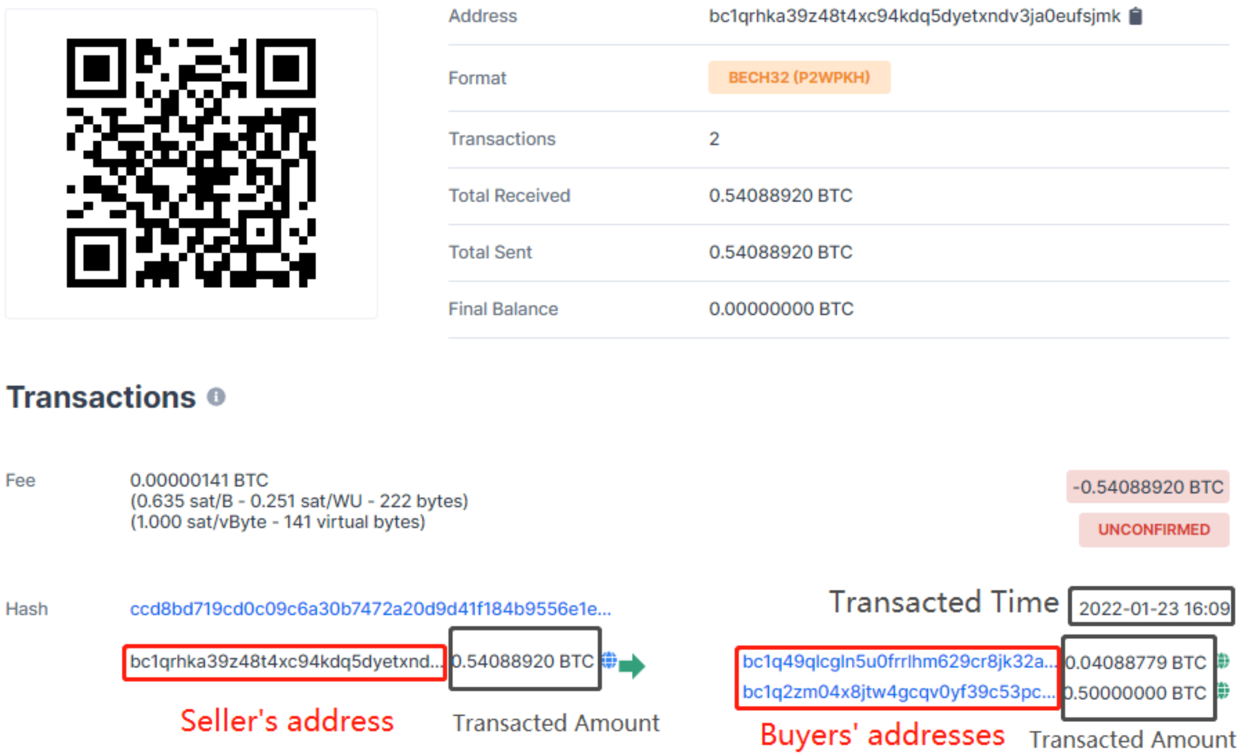


Figure 3: The Role of Priors on Sentiment Contagion

This figure shows how users with different priors change their sentiment in response to conversations. The black solid line represents the users with positive priors. The blue dash line represents the users with neutral priors. The red dot line represents the users with negative priors. For the purpose of comparison on the slopes, we ignore the intercept. Moreover, we use the Seemingly Unrelated Regression to test the difference in slope for different priors. The difference between slopes of positive and negative priors is significant at 1%.

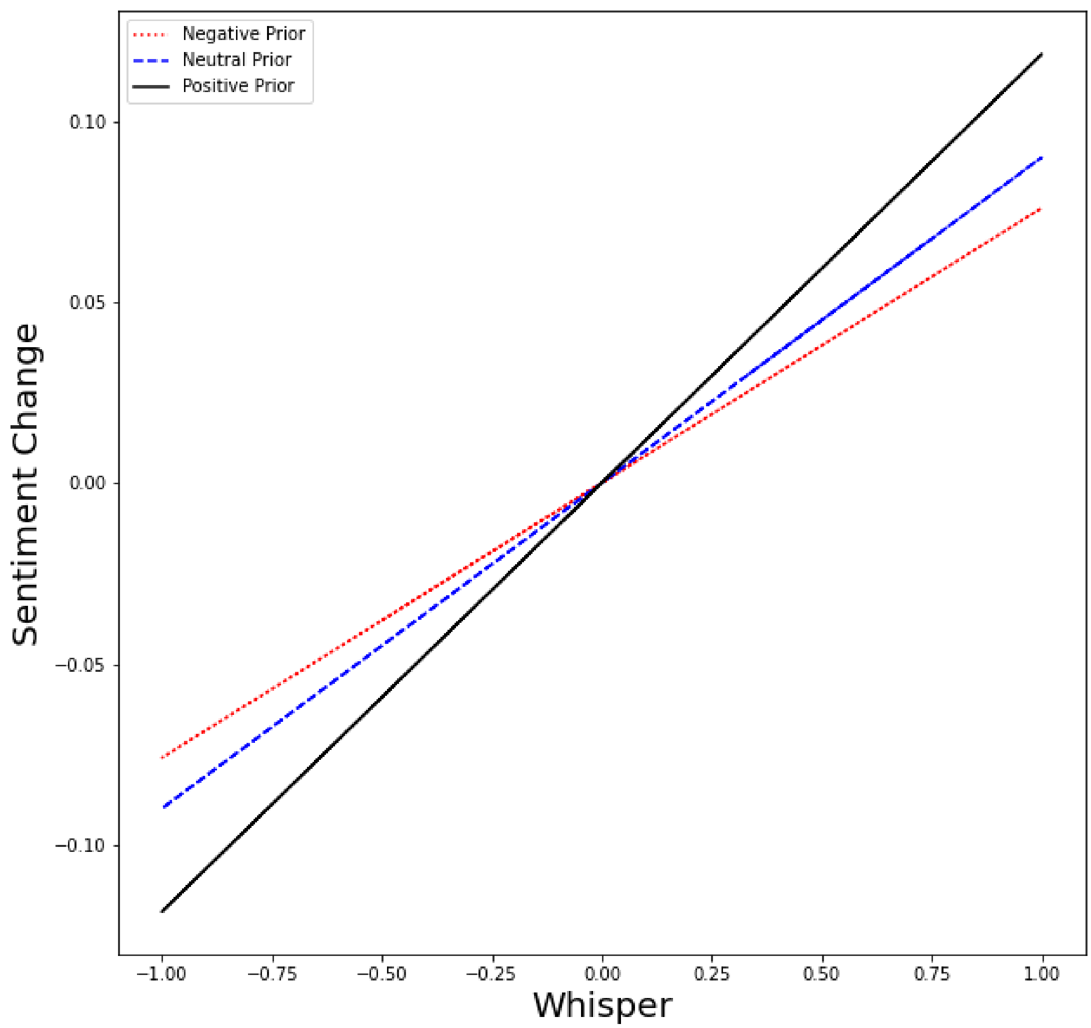
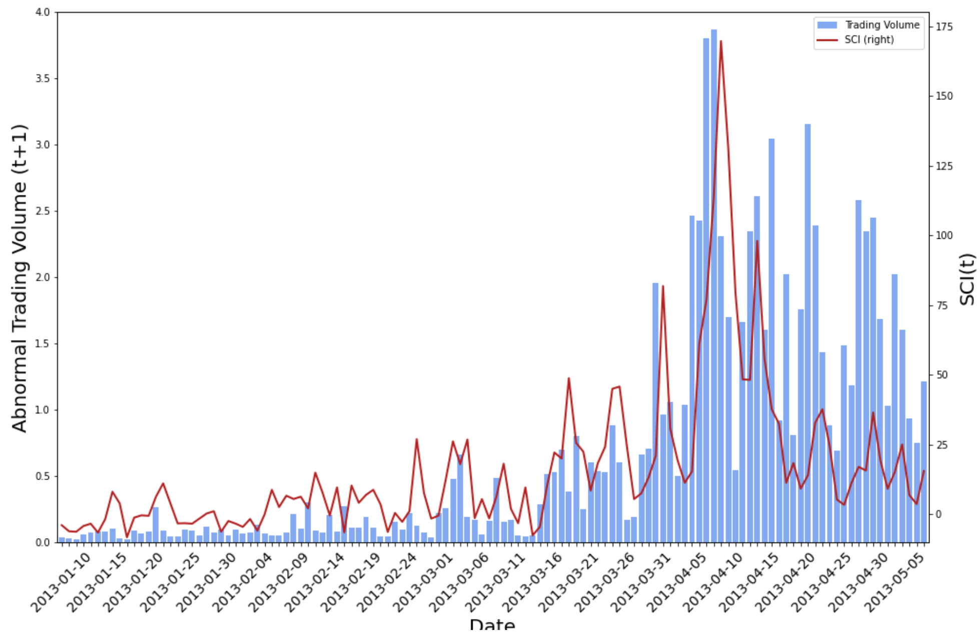


Figure 4: Sentiment Contagion and Trading Volume in Bubbles

This figure plots the sentiment contagion indicator (SCI) and the trading volume in the next day during the two identified bubble formation episodes. The SCI indicator is defined as the total number of affected investors within a day, and we remove the trend by regressing it on the weekday and year-month-pair dummies. The volume is the raw trading volume in the next day.



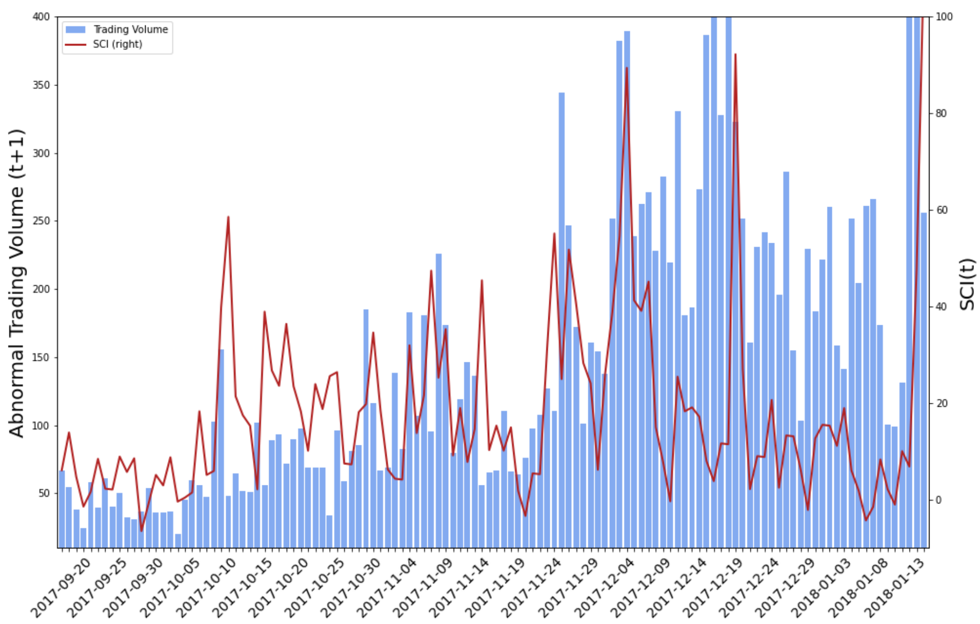


Table 1: Summary Statistics

This table tabulates the summary statistics of the key variables. In Panel A, we present features of investor sentiment in posts aggregated at the user level. To obtain the summary statistics for the average sentiment, we first calculate the mean sentiment of posts published by each user, then report the corresponding statistics based on the whole user population. In Panel B, we present features of investor sentiment in posts aggregated at the daily level. In Panel C, we present features of investor sentiment in posts aggregated at the conversation level. In Panel D, we present summary statistics of market variables, RavenPack news sentiment and google search volume. Daily Return is the annualized daily return of Bitcoin, Return Volatility (Within Day) is the within-day volatility of Bitcoin based on the hourly Bitcoin return. Number of Transactions is the total number of Bitcoin transactions within one day. Number of Bitcoins Traded is the total number of Bitcoins being traded within one day. Total Dollar Volume (in millions) is the total trading volume of Bitcoin measured in millions of dollars within one day. RavenPack News Sentiment is the sentiment score for the news on Bitcoin in the RavenPack database. We normalize the sentiment score to be between $[-1, 1]$. Abnormal Google Search Volume is defined as the difference between the Google Search Volume Index for the keyword "Bitcoin" and its past one-month mean, divided by the lagged one-month mean. Our sample spans from Jan 5, 2012 to July 30, 2018.

	count	mean	p50	sd	min	max
Panel A: Post Activities at User level						
Average Sentiment	37262	0.329	0.343	0.517	-1.000	1.000
Standard Deviation of Sentiment	23458	0.604	0.635	0.301	0.000	1.414
Number of Posts	37262	11.582	2.000	48.697	1.000	2778.000
Panel B: Post Activities at Daily level						
Average Sentiment	2275	0.268	0.274	0.123	-0.333	0.833
Standard Deviation of Sentiment	2274	0.680	0.680	0.051	0.392	0.917
Number of Users	2275	128.805	83.000	134.797	1.000	986.000
Number of Posts	2275	189.699	134.000	176.789	1.000	1411.000
Panel C: Post Activities at Conversation level						
Average Sentiment	15729	0.228	0.242	0.343	-1.000	1.000
Standard Deviation of Sentiment	14187	0.648	0.666	0.202	0.000	1.414
Number of Users	15729	20.968	9.000	51.329	1.000	2422.000
Number of Posts	15729	27.438	11.000	125.493	1.000	11031.000
Median Gap in Days between Consecutive Posts	14215	0.572	0.027	16.078	0.000	1129.115
Panel D: Other Data Sources						
Daily Return(Annualized)	2281	1.602	0.887	17.273	-177.093	146.518
Return Volatility(Within Day)	2281	0.040	0.030	0.041	0.000	0.599
Number of Transactions	2281	11710.698	5919.000	16032.126	0.000	181616.000
Number of Bitcoins Traded	2281	11157.683	8232.481	11166.875	0.000	137070.178
Total Dollar Volume (in millions)	2281	23.128	3.438	54.333	0.000	779.448
RavenPack News Sentiment	676	0.159	0.440	0.551	-0.660	0.660
Abnormal Google Search Volume	2209	0.018	-0.049	0.330	-0.616	2.813

Table 2: Correlation Matrix

This table tabulates the correlation matrix between investor sentiment and other key variables in this paper. For the first four variables— Average Sentiment, Standard Deviation of Sentiment, Number of Users and Number of Posts, we follow the procedure in Panel B of Table 1to get the aggregated post activities at the daily level. Daily Return is the annualized daily return of Bitcoin, Return Volatility (Within Day) is the within-day volatility of Bitcoin based on the hourly Bitcoin return. Number of Transactions is the total number of Bitcoin transactions within one day. Number of Bitcoins Traded is the total number of Bitcoins being traded within one day. Total Dollar Volume (in millions) is the total trading volume of Bitcoin measured in millions of dollars within one day. RavenPack News Sentiment is the sentiment score for the news on Bitcoin in the RavenPack database. We normalize the sentiment score to be between $[-1, 1]$. Abnormal Google Search Volume is defined as the difference between the Google Search Volume Index for the keyword "Bitcoin" and its past one-month mean, divided by the lagged one-month mean. Our sample spans from Jan 5, 2012 to July 30, 2018.

Variables	Average Sentiment	Standard Deviation of Sentiment	Number of Users	Number of Posts	Daily Return (Annualized)	Return Volatility (Within Day)	Number of Transactions	Number of Bitcoins Traded	Total Dollar Volume (in millions)	RavenPack News Sentiment	Abnormal Google Search Volume
Average Sentiment	1.000										
Standard Deviation of Sentiment	-0.492	1.000									
	(0.000)										
Number of Users	0.284	-0.123	1.000								
	(0.000)	(0.000)									
Number of Posts	0.239	-0.075	0.975	1.000							
	(0.000)	(0.000)	(0.000)								
Daily Return(Annualized)	0.124	-0.076	-0.013	-0.025	1.000						
	(0.000)	(0.000)	(0.544)	(0.227)							
Return Volatility(Within Day)	-0.183	0.138	0.264	0.381	-0.141	1.000					
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)						
Number of Transactions	0.183	-0.121	0.767	0.738	-0.046	0.338	1.000				
	(0.000)	(0.000)	(0.000)	(0.000)	(0.027)	(0.000)					
Number of Bitcoins Traded	-0.128	0.134	0.280	0.367	-0.068	0.581	0.448	1.000			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)				
Total Dollar Volume (in millions)	0.211	-0.144	0.738	0.694	-0.041	0.292	0.946	0.316	1.000		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.052)	(0.000)	(0.000)	(0.000)			
RavenPack News Sentiment	0.257	-0.202	-0.223	-0.231	0.339	-0.238	-0.257	-0.225	-0.228	1.000	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
Abnormal Google Search Volume	-0.019	0.078	0.103	0.184	0.067	0.402	0.091	0.295	0.088	0.080	1.000
	(0.380)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.040)	

Table 3: Contagion Effect

This table presents panel regression analysis of sentiment change on Whisper sentiment. The dependent variable is the sentiment change defined as the revision in sentiment between a user's two consecutive posts post[0] and post[1]. We are primarily interested in the case when post[0] and post[1] are published within a 24-hours window, and post[0] and post[1] can be in two different conversations. The main independent variable of interest is the Whisper sentiment variable defined as the average level of other users' sentiment in the same threads of post[0] and post[1], and published between the timestamps of post[0] and post[1]. We calculate the average levels of sentiment in RavenPack news as controls for news arrivals between two consecutive posts, 24 hours before post[0] and 48 hours before post[0]. We control for market information by controlling for three Bitcoin market variables: Bitcoin Return, Bitcoin Volatility and Number of Transactions. Each of three variable is the cumulative sum between the timestamps of post[0] and post[1]. We calculate forum sentiment as the average sentiment of posts on the Bitcointalk forum published between timestamps of post[0] and post[1]. In all columns we add user and date fixed effect. We also control for user priors. The t-statistics (in parentheses) are based on standard errors clustered by user and date.

	(1)	(2)	(3)	(4)
Whisper sentiment	0.030***	0.030***	0.036***	0.035***
	(5.61)	(5.60)	(5.91)	(5.69)
RavenPack News Sentiment		0.012	0.005	0.004
between Post[0] and Post[1]		(1.17)	(0.41)	(0.40)
RavenPack News Sentiment		0.001	0.010	0.010
24 hours before Post[0]		(0.10)	(0.83)	(0.82)
RavenPack News Sentiment		-0.006	-0.006	-0.007
48 hours before Post[0]		(-0.53)	(-0.54)	(-0.59)
Bitcoin Return			0.239***	0.243***
			(4.27)	(4.36)
Bitcoin Volatility			0.213	0.186
			(1.28)	(1.10)
Number of Transactions			-0.000*	-0.000*
			(-1.66)	(-1.67)
Forum Sentiment				-0.053***
				(-3.93)
Adjusted R-Squared	0.345	0.345	0.346	0.346
N	124,970	124,970	105,552	103,682
User FE	YES	YES	YES	YES
Date FE	YES	YES	YES	YES
Control For Prior	YES	YES	YES	YES

Table 4: Placebo Tests

This table presents the placebo test for the contagion effect. The dependent variable is the sentiment change defined as the revision in sentiment between a user's two consecutive posts post[0] and post[1]. We are primarily interested in the case when post[0] and post[1] are published within a 24-hours window, and post[0] and post[1] can be in two different conversations. The independent variable is the average level of sentiment in a random conversation that the user did not participate that happens between the timestamps of post[0] and post[1]. The control variables are the same as those in Table 3. In all columns we add user and date fixed effect. We also control for user priors. The t-statistics (in parentheses) are based on standard errors clustered by user and date.

	(1)	(2)	(3)	(4)
Sentiment in Random Conversation	-0.003	-0.003	-0.008*	-0.001
	(-0.91)	(-0.92)	(-1.87)	(-0.28)
RavenPack News Sentiment		0.007	0.004	0.004
between Post[0] and Post[1]		(0.76)	(0.38)	(0.43)
RavenPack News Sentiment		-0.007	0.012	0.013
24 hours before Post[0]		(-0.77)	(1.08)	(1.11)
RavenPack News Sentiment		-0.010	-0.007	-0.007
48 hours before Post[0]		(-0.99)	(-0.64)	(-0.65)
Bitcoin Return			0.257***	0.261***
			(4.84)	(4.90)
Bitcoin Volatility			0.151	0.135
			(1.04)	(0.93)
Number of Transactions			-0.000**	-0.000**
			(-2.11)	(-2.13)
Forum Sentiment				-0.049***
				(-4.05)
Adjusted R-Squared	0.346	0.346	0.346	0.346
N	153,292	153,292	118,772	118,772
User FE	YES	YES	YES	YES
Date FE	YES	YES	YES	YES
Control For Prior	YES	YES	YES	YES

Table 5: Contagion Effect by User Feature

This table presents the contagion effect for users with different features. The dependent variable is the sentiment change defined as the revision in sentiment between a user's two consecutive posts post[0] and post[1]. We are primarily interested in the case when post[0] and post[1] are published within a 24-hours window, and post[0] and post[1] can be in two different conversations. The main independent variable of interest is the Whisper sentiment variable defined as the average level of other users' sentiment in the same threads of post[0] and post[1], and published between the timestamps of post[0] and post[1]. We interact different features with the independent variable. In column (1), we define naive users are the users with low ranks (non-legendary rankings) on the Bitcointalk website. In column (2), female users are the ones who voluntarily reveal them as female. In column (3), we define users whose mother tougue is English as those who come from one of the following countries: Australia, Canada, New Zealand, the United Kingdom and the United States of America. In column (4), we define old users as those who are above the age of 40. The control variables are the same as those in Table 3. In all columns we add user and date fixed effect. We also control for user priors. The t-statistics (in parentheses) are based on standard errors clustered by user and date.

	(1) Naive User	(2) Female User	(3) English as Mother Tongue	(4) Old User
User Feature * Whisper sentiment	0.041***	-0.093*	-0.155**	0.033
	(3.13)	(-1.90)	(-2.43)	(1.59)
Adjusted R-Squared	0.346	0.351	0.334	0.346
N	103,682	23,364	8,919	103,682
User FE	YES	YES	YES	YES
Date FE	YES	YES	YES	YES
Control For Prior	YES	YES	YES	YES
Other Controls	YES	YES	YES	YES

Table 6: Contagion Effect by Conversation Feature

This table presents the contagion effect for conversations with different features. The dependent variable is the sentiment change defined as the revision in sentiment between a user's two consecutive posts post[0] and post[1]. We are primarily interested in the case when post[0] and post[1] are published within a 24-hours window, and post[0] and post[1] can be in two different conversations. The main independent variable of interest is the Whisper sentiment variable defined as the average level of other users' sentiment in the same threads of post[0] and post[1], and published between the timestamps of post[0] and post[1]. We interact different features with the independent variable. In column (1), we include a dummy variable for conversations in which at least one sophisticated user participates. Sophisticated users are the users with high ranks (legendary ranking) on the Bitcointalk website. In column (2), we include the number of posts in a conversation. In column (3), we consider the feature of Attitude dispersion, calculated as the standard deviation of sentiment in a conversation. In column (4), we consider both the number of posts and the standard deviation of sentiment in a conversation. In column (3) and (4), we require the number of posts in a conversation to be larger than one, in order to get a reasonable value of standard deviation of sentiment. The control variables are the same as those in Table 3. In all columns we add user and date fixed effect. All coefficients are multiplied by 100. We also control for user priors. The t-statistics (in parentheses) are based on standard errors clustered by user and date.

	(1) Sophisticated Users	(2) Number of Posts	(3) Attitude Dispersion	(4) Number of Posts& Attitude Dispersion
Sophisticated Users* Whisper sentiment	0.055**			
	(2.00)			
Number of Posts* Whisper sentiment		0.004***		0.004***
		(4.57)		(4.18)
Attitude Dispersion* Whisper sentiment			-0.177***	-0.192***
			(-2.93)	(-3.19)
Adjusted R-Squared	0.346	0.346	0.348	0.348
N	103,682	103,682	87,988	87,988
User FE	YES	YES	YES	YES
Date FE	YES	YES	YES	YES
Control For Prior	YES	YES	YES	YES
Controls	YES	YES	YES	YES

Table 7: Real Effect

This table presents regression analysis of users' subsequent trading decision on the sentiment change between their Post[0] and Post[1]. In column (1) and (2), the dependent variable is a dummy that equals one if the user's first transaction after publishing post[1] is to buy bitcoins and if this first transaction occurs within a 12-hour window after post [1], and zero otherwise. In column (3) and (4), the dependent variable is a dummy that equals one if the user's first transaction is to buy bitcoins and zero otherwise. Compared to column (1) and (2), in column (3) and (4), we do not restrict a time window for the first transaction after post [1]. The control variables are the same as those in Table 3. In all columns we add user and date fixed effects. We multiply each coefficient by 100. The t-statistics (in parentheses) are calculated based on standard errors clustered by user and date.

	(1) First transaction as buy within 12 hours	(2) First transaction as buy within 12 hours	(3) First transaction as buy	(4) First transaction as buy
Sentiment Change	0.202**	0.217**	0.210	0.382*
between Post[0] and Post[1]	(2.43)	(1.99)	(1.44)	(1.72)
RavenPack News Sentiment		-0.323		1.248
between Post[0] and Post[1]		(-1.03)		(1.33)
RavenPack News Sentiment		-0.355		-0.999
24 hours before Post[0]		(-0.75)		(-1.16)
RavenPack News Sentiment		-0.582*		0.302
48 hours before Post[0]		(-1.72)		(0.34)
Bitcoin Return		0.713		4.495
		(0.31)		(0.67)
Bitcoin Volatility		-7.490		-41.627***
		(-1.06)		(-2.94)
Number of Transactions		-0.000		0.000***
		(-0.50)		(2.93)
Forum Sentiment		-0.089		-0.676
		(-0.21)		(-1.01)
Adjusted R-Squared	0.147	0.106	0.381	0.349
N	50,523	34,061	50,523	34,061
Fund FE	YES	YES	YES	YES
Date FE	YES	YES	YES	YES
Control for Prior	YES	YES	YES	YES

Table 8: Sentiment Contagion Indicator and Market Dynamics in the Future

This table presents the predictive power of sentiment contagion indicator (SCI) on future market dynamics. The main explanatory variable of interest is the SCI indicator, which is defined as the daily number of users who revise their sentiment after conversations towards the direction of the average sentiment in conversation. We eliminate the date and weekday fixe effects for SCI. In column (1) and (2), the main dependent variable of interest is the future trading volume of Bitcoin in the next period. We normalize the trading volume using its past two-week mean following Da, Engelberg, and Gao (2011). In column (3) and (4), the main dependent variable of interest is the future volatility of Bitcoin return in the next period. In column (5) and (6), the main dependent variable of interest is the future Bitcoin return(in log) in the next period . We calculate the average levels of sentiment in RavenPack news in the past 14 days as controls for news arrivals. We control for market information by controlling for two Bitcoin market variables: Bitcoin Volatility and Number of Transactions. Each of two variable is the cumulative sum in the past 14 days. We calculate forum average sentiment as the average sentiment of posts on the Bitcointalk forum published in the past 14 days. We multiply each coefficient by 100. The t-statistics (in parentheses) are based on Newey-West corrected standard errors using a lag of 14 days . *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

	(1) Trading Volume	(2) Trading Volume	(3) Return Volatility	(4) Return Volatility	(5) Log Return	(6) Log Return
SCI	22.804**	21.550**	2.734***	2.007***	0.541	0.017
	(2.19)	(2.30)	(5.05)	(3.37)	(0.12)	(0.00)
Bitcoin Volatility		-2.5e+03***		179.079***		504.842
in Past 14 Days		(-3.23)		(5.66)		(1.47)
Total Dollar Volume		7270.324		12.005		-865.266
in Past 14 Days(in millions)		(1.44)		(0.13)		(-0.81)
RavenPack News Sentiment		538.933**		2.732		126.037
in Past 14 Days		(2.06)		(0.38)		(1.07)
Forum average sentiment		-977.152		-107.816***		309.659
in Past 14 Days		(-1.45)		(-5.48)		(0.81)
Adjusted R-Squared	0.016	0.045	0.229	0.335	0.000	0.002
N	2,418	2,412	2,467	2,459	2,467	2,459
User FE	NO	NO	NO	NO	NO	NO
Date FE	NO	NO	NO	NO	NO	NO
Controls	NO	YES	NO	YES	NO	YES

Table 9: Summary Statistics: Bubble vs Non-Bubble Episodes

This table tabulates the summary statistics of the key features in bubble and non-bubble episodes, and the difference of the key variables in bubble and non-bubble episodes. A bubble is identified in two steps. First, we identify the days when there are at least 100% increase OF Bitcoin returns during the past one month and 40% decline in Bitcoin value within the subsequent month. If we get multiple days, we merge them to get a longer window. We pick the day with highest Bitcoin price from this merged window of days as the peak day of each bubble. The bubble formation episode is then defined as a continuous window of days before its peak day, and we require that the past 30-day return of each day in this window be higher than 30%. The bubble burst episode is defined as the one month window right after the peak day, with a return of less than 40%. In Panel A, we describe market variables. In Panel B, we describe social interactions and investor sentiment. We rely on seemingly unrelated regression (SUR) to test the joint significance of differences in all features. We report the joint F-statistic and its corresponding p-value. Our sample spans from Jan 5, 2012 to July 30, 2018.

	Bubble Formation		No Bubble		Bubble Formation minus No Bubble	
Features	Mean	Standard deviation	Mean	Standard deviation	Difference	t-statistic
Panel A: Market Variables						
Daily Return(Annualized)	9.506	18.84	1.037	15.45	8.47	6.91
Return Volatility(Within Day)	0.056	0.05	0.036	0.03	0.02	7.61
Number of Transactions	19293.889	25823.28	10075.882	11887.79	9218.007	8.72
Total Dollar Volume (in millions)	58.306	83.70	16.551	37.52	41.755	12.42
RavenPack News Sentiment	0.301	0.47	0.149	0.56	0.152	2.34
Abnormal Google Search Volume	0.324	0.48	-0.007	0.28	0.331	13.97
Panel B: Panel B: Social Interactions and Investor Sentiment						
Average Sentiment	0.292	0.12	0.265	0.12	0.027	2.86
Standard Deviation of Sentiment	0.671	0.05	0.681	0.05	-0.01	-2.57
Number of Users	205.228	161.63	116.408	123.44	88.819	8.99
Fraction of Sophisticated Users	0.901	0.05	0.915	0.05	-0.014	-3.36
Number of Posts	290.239	208.69	172.542	160.32	117.697	9.18
Total Number of Positive Posts	164.494	131.67	90.565	90.87	73.929	10.01
Fraction of Posts with Positive Sentiment	0.526	0.09	0.501	0.09	0.024	3.45
Joint F-statistic						17.41
p-value (Probability>F)						0.00

Table 10: Post Decisions and Recent Bitcoin Returns

This table presents the association between users' post decisions and Bitcoin returns in the past 14 days. In column (1) and (2), the dependent variable $PostDecisions_{i,t}$ is a dummy variable that equals one if user i makes at least one post on day t and zero otherwise. We focus on users' active days after their registration. In column (3) to (6), we investigate how, conditional on users' post decisions, recent Bitcoin returns affect sentiment in posts . In column (3) and (4), the dependent variable $PositivePosts_{i,t}$ equals to one if user i publishes a positive posts on day t and zero otherwise. In column (5) and (6), the dependent variable $PessimisticPosts_{i,t}$ equals to one if user i publishes a negative posts on day t and zero otherwise. We calculate the average levels of sentiment in RavenPack news in the past 14 days as controls for news arrivals. We control for market information by controlling for two Bitcoin market variables: Bitcoin Volatility and Number of Transactions. Each of two variable is the cumulative sum in the past 14 days. We calculate forum average sentiment as the average sentiment of posts on the Bitcointalk forum published in the past 14 days. In all columns we add date fixed effect. We multiply each coefficient by 100. The t-statistics (in parentheses) are based on standard errors clustered by user and date. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

	(1) Post Decisions	(2) Post Decisions	(3) Positive Posts	(4) Positive Posts	(5) Pessimistic Posts	(6) Pessimistic Posts
Bitcoin Return	0.166***	0.254***	5.807***	1.986***	-4.274***	-1.789***
in Past 14 Days	(3.96)	(6.08)	(9.25)	(3.90)	(-8.46)	(-4.11)
Bitcoin Volatility		0.436***		2.035**	in Past 14 Days	-1.203
		(6.47)		(2.41)		(-1.63)
Number of Transactions		0.002***		-0.032***		0.041***
in Past 14 Days		(3.65)		(-5.22)		(8.65)
RavenPack News Sentiment		-0.042***		2.279***		-1.695***
in Past 14 Days		(-2.68)		(7.51)		(-7.81)
Forum average sentiment		-0.872***		35.007***		-21.679***
in Past 14 Days		(-10.02)		(15.69)		(-12.58)
Adjusted R-Squared	0.073	0.074	0.042	0.044	0.021	0.022
N	60147616	60034776	436,711	436,644	436,711	436,644
User FE	YES	YES	YES	YES	YES	YES
Date FE	NO	NO	NO	NO	NO	NO
Controls	YES	YES	YES	YES	YES	YES

Table 11: Contagion Effect in Bubble Formation Episodes

This table presents the contagion effect for bubble formation episodes. The dependent variable is the sentiment change defined as the revision in sentiment between a user's two consecutive posts post[0] and post[1]. We are primarily interested in the case when post[0] and post[1] are published within a 24-hours window, and post[0] and post[1] can be in two different conversations. The main independent variable of interest is the Whisper sentiment variable defined as the average level of other users' sentiment in the same threads of post[0] and post[1], and published between the timestamps of post[0] and post[1]. We create two types of dummy variables for bubble formation episodes. In column (1), the dummy variable equals to zero if it is in the non-bubble episodes. In column (2), the dummy variable equals to zero if it is either in the non-bubble episodes or the bubble burst episodes. The control variables are the same as those in Table 3. In all columns we add user and date fixed effect. We also control for user priors. The t-statistics (in parentheses) are based on standard errors clustered by user and date.

	(1) Bubble Formation vs Non-Bubble	(2) Bubble Formation vs Non-Bubble and Bubble Burst
Bubble Formation* Whisper sentiment	0.030*	0.029*
	(1.91)	(1.88)
Whisper sentiment	0.052***	0.053***
	(8.28)	(8.76)
Adjusted R-Squared	0.347	0.346
N	95,519	102,230
User FE	YES	YES
Date FE	YES	YES
Control For Prior	YES	YES
Other Controls	YES	YES

References

Antweiler, W. and 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.

Baker, M. and Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The journal of Finance*, 61(4):1645–1680.

Baker, M. and Wurgler, J. (2013). Behavioral corporate finance: An updated survey. volume 2, chapter Chapter 5, pages 357–424. Elsevier.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Barber, B. and Odean, T. (2013). The behavior of individual investors. volume 2, chapter Chapter 22, pages 1533–1570. Elsevier.

Barber, B. M. and Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, 116(1):261–292.

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

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

Brown, B., Bakken, J., Ameringer, S., and Mahon, S. (2008). A comprehensive conceptualization of the peer influence process in adolescence. *Understanding Peer Influence In Children and Adolescents*, pages 17–44.

Brown, G. and Cliff, M. (2005). Investor sentiment and asset valuation. *The Journal of Business*, 78(2):405–440.

Da, Z., Engelberg, J., and Gao, P. (2011). In search of attention. *The Journal of Finance*, 66(5):1461–1499.

Da, Z., Engelberg, J., and Gao, P. (2014). The sum of all fears investor sentiment and asset prices. *The Review of Financial Studies*, 28(1):1–32.

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, 98(4):703–738.

DeFusco, A. A., Nathanson, C. G., and Zwick, E. (2017). Speculative dynamics of prices and volume. Technical report, National Bureau of Economic Research.

D'Acunto, F., Hoang, D., Paloviita, M., and Weber, M. (2019). Iq, expectations, and choice. Technical report, National Bureau of Economic Research.

Feng, L. and Seasholes, M. S. (2004). Correlated trading and location. *The Journal of Finance*, 59(5):2117–2144.

Giglio, S., Maggiori, M., Stroebel, J., and Utkus, S. (2020). Inside the mind of a stock market crash. Working Paper 27272, National Bureau of Economic Research.

Giglio, S., Maggiori, M., Stroebel, J., and Utkus, S. (2021). Five facts about beliefs and portfolios. *American Economic Review*, 111(5):1481–1522.

Greenwood, R. and Shleifer, A. (2014). Expectations of returns and expected returns. *The Review of Financial Studies*, 27(3):714–746.

Griffin, J. M. and Shams, A. (2020). Is bitcoin really untethered? *The Journal of Finance*, 75(4):1913–1964.

Grinblatt, M., Keloharju, M., and Linnainmaa, J. (2011). Iq and stock market participation. *The Journal of Finance*, 66(6):2121–2164.

Han, B., Hirshleifer, D., and Walden, J. (2022). Social transmission bias and investor behavior. *Journal of Financial and Quantitative Analysis*, 57(1):390–412.

Hasher, L., Goldstein, D., and Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of verbal learning and verbal behavior*, 16(1):107–112.

Heimer, R. Z. (2016). Peer pressure: Social interaction and the disposition effect. *The Review of Financial Studies*, 29(11):3177–3209.

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

1831.

Hong, H., Kubik, J. D., and Stein, J. C. (2004). Social interaction and stock-market participation. *The Journal of Finance*, 59(1):137–163.

Hong, H., Kubrik, J. D., and Stein, J. C. (2005). Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. *The Journal of Finance*, 60(6):2801–2824.

Huang, S., Hwang, B.-H., and Lou, D. (2021). The rate of communication. *Journal of Financial Economics*, 141(2):533–550.

Kuchler, T. and Zafar, B. (2019). Personal experiences and expectations about aggregate outcomes. *The Journal of Finance*, 74(5):2491–2542.

Lee, C. M. C., Shleifer, A., and Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. *Journal of Finance*, 46(1):75–109.

Lemmon, M. and Portniaguina, E. (2006). Consumer confidence and asset prices: Some empirical evidence. *Review of Financial Studies*, 19(4):1499–1529.

Liao, J., Peng, C., and Zhu, N. (2021). Extrapolative bubbles and trading volume. *Review of Financial Studies*.

Lundeberg, M. A., Fox, P. W., and Punćcohaŕ, J. (1994). Highly confident, but wrong: Gender differences and similarities in confidence judgments. *Journal of Educational Psychology*, 86:114–121.

Qiu, L. and Welch, I. (2004). Investor sentiment measures. NBER Working Papers 10794, National Bureau of Economic Research, Inc.

Shiller, R. J. (1989). Comovements in stock prices and comovements in dividends. *The Journal of Finance*, 44(3):719–729.

Shiller, R. J. (2015). *Irrational exuberance*. Princeton university press.

Shiller, R. J. (2017). Narrative economics. *American economic review*, 107(4):967–1004.

Shiller, R. J. and Pound, J. (1989). Survey evidence on diffusion of interest and information among investors. *Journal of Economic Behavior & Organization*, 12(1):47–66.

Soo, C. K. (2018). Quantifying sentiment with news media across local housing markets. *The Review of Financial Studies*, 31(10):3689–3719.

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.

Table A1: Contagion Effect: Same Conversation

This table presents panel regression analysis of sentiment change on Whisper sentiment. The dependent variable is the sentiment change defined as the revision in sentiment between a user's two consecutive posts post[0] and post[1]. We are primarily interested in the case when post[0] and post[1] are published within a 24-hours window, and we restrict post[0] and post[1] to be in the same conversations. The main independent variable of interest is the Whisper sentiment variable defined as the average level of other users' sentiment in the same thread of post[0] and post[1], and published between the timestamps of post[0] and post[1]. We calculate the average levels of sentiment in RavenPack news as controls for news arrivals between two consecutive posts, 24 hours before post[0] and 48 hours before post[0]. We control for market information by controlling for three Bitcoin market variables: Bitcoin Return, Bitcoin Volatility and Number of Transactions. Each of three variable is the cumulative sum between the timestamps of post[0] and post[1]. We calculate forum sentiment as the average sentiment of posts on the Bitcointalk forum published between timestamps of post[0] and post[1]. In all columns we add user and date fixed effect. We also control for user priors. The t-statistics (in parentheses) are based on standard errors clustered by user and date.

	(1)	(2)	(3)	(4)
Whisper	0.029***	0.029***	0.028**	0.030**
	(2.90)	(2.90)	(2.33)	(2.40)
RavenPack News Sentiment		0.002	0.003	-0.005
between Post[0] and Post[1]		(0.06)	(0.09)	(-0.14)
RavenPack News Sentiment		0.033	0.017	0.018
24 hours before Post[0]		(1.05)	(0.47)	(0.47)
RavenPack News Sentiment		0.026	0.028	0.024
48 hours before Post[0]		(0.92)	(0.89)	(0.73)
Bitcoin Return			0.109	0.104
			(0.98)	(0.93)
Bitcoin Volatility			-0.002	-0.037
			(-0.01)	(-0.10)
Number of Transactions			0.000	0.000
			(0.22)	(0.07)
Forum Sentiment				-0.041*
				(-1.68)
Adjusted R-Squared	0.325	0.324	0.330	0.331
N	29,616	29,616	23,117	22,246
User FE	YES	YES	YES	YES
Date FE	YES	YES	YES	YES
Control For Prior	YES	YES	YES	YES

Table A2: Contagion Effect: 12 Hours as the Interval

This table presents panel regression analysis of sentiment change on Whisper sentiment using the 12 hours as the interval. The dependent variable is the sentiment change defined as the revision in sentiment between a user's two consecutive posts post[0] and post[1]. We are primarily interested in the case when post[0] and post[1] are published within a 12-hours window. The main independent variable of interest is the Whisper sentiment variable defined as the average level of other users' sentiment in the same thread of post[0] and post[1], and published between the timestamps of post[0] and post[1]. We calculate the average levels of sentiment in RavenPack news as controls for news arrivals between two consecutive posts, 24 hours before post[0] and 48 hours before post[0]. We control for market information by controlling for three Bitcoin market variables: Bitcoin Return, Bitcoin Volatility and Number of Transactions. Each of three variable is the cumulative sum between the timestamps of post[0] and post[1]. We calculate forum sentiment as the average sentiment of posts on the Bitcointalk forum published between timestamps of post[0] and post[1]. In all columns we add user and date fixed effect. We also control for user priors. The t-statistics (in parentheses) are based on standard errors clustered by user and date.

	(1)	(2)	(3)	(4)
Whisper	0.024***	0.024***	0.030***	0.029***
	(4.14)	(4.12)	(4.24)	(3.97)
RavenPack News Sentiment		0.023	0.009	0.009
between Post[0] and Post[1]		(1.62)	(0.59)	(0.56)
RavenPack News Sentiment		0.008	0.010	0.010
24 hours before Post[0]		(0.50)	(0.59)	(0.55)
RavenPack News Sentiment		-0.003	-0.003	-0.005
48 hours before Post[0]		(-0.20)	(-0.15)	(-0.25)
Bitcoin Return			0.226***	0.234***
			(3.03)	(3.13)
Bitcoin Volatility			0.139	0.120
			(0.71)	(0.60)
Number of Transactions			-0.000	-0.000
			(-0.71)	(-0.81)
Forum Sentiment				-0.044***
				(-2.97)
Adjusted R-Squared	0.342	0.342	0.344	0.343
N	83,694	83,694	64,497	62,635
User FE	YES	YES	YES	YES
Date FE	YES	YES	YES	YES
Control For Prior	YES	YES	YES	YES

Table A3: Contagion Effect: More Controls for News Arrivals

This table presents panel regression analysis of sentiment change on Whisper sentiment with more controls. The dependent variable is the sentiment change defined as the revision in sentiment between a user's two consecutive posts post[0] and post[1]. We are primarily interested in the case when post[0] and post[1] are published within a 24-hours window. The main independent variable of interest is the Whisper sentiment variable defined as the average level of other users' sentiment in the same thread of post[0] and post[1], and published between the timestamps of post[0] and post[1]. We calculate the average levels of sentiment in RavenPack news as controls for news arrivals between two consecutive posts, and past arrivals up to past 7 days in the past. We control for market information by controlling for three Bitcoin market variables: Bitcoin Return, Bitcoin Volatility and Number of Transactions. Each of three variable is the cumulative sum between the timestamps of post[0] and post[1]. We calculate forum sentiment as the average sentiment of posts on the Bitcointalk forum published between timestamps of post[0] and post[1]. In all columns we add user and date fixed effect. We also control for user priors. The t-statistics (in parentheses) are based on standard errors clustered by user and date.

	(1)	(2)	(3)	(4)
Whisper	0.030***	0.030***	0.036***	0.035***
	(5.61)	(5.61)	(5.92)	(5.69)
RavenPack News Sentiment		0.013	0.004	0.004
between Post[0] and Post[1]		(1.24)	(0.40)	(0.40)
RavenPack News Sentiment		0.002	0.010	0.010
24 hours before Post[0]		(0.20)	(0.79)	(0.81)
RavenPack News Sentiment between		-0.004	-0.007	-0.007
48 and 24 hours before Post[0]		(-0.31)	(-0.53)	(-0.53)
RavenPack News Sentiment		0.012	0.006	0.007
between 72 and 48 hours before Post[0]		(0.92)	(0.42)	(0.50)
RavenPack News Sentiment		-0.011	-0.014	-0.014
between 96 and 72 hours before Post[0]		(-0.87)	(-1.11)	(-1.05)
RavenPack News Sentiment		-0.000	0.000	0.002
between 120 and 96 hours before Post[0]		(-0.02)	(0.04)	(0.19)
RavenPack News Sentiment		-0.011	-0.009	-0.011
between 144 and 120 hours before Post[0]		(-0.98)	(-0.67)	(-0.83)
Bitcoin Return			0.241***	0.245***
			(4.27)	(4.35)
Bitcoin Volatility			0.206	0.181
			(1.25)	(1.08)
Number of Transactions			-0.000*	-0.000*
			(-1.68)	(-1.70)
Forum Sentiment				-0.053***
				(-3.94)
Adjusted R-Squared	0.345	0.345	0.346	0.346
N	125,065	125,065	105,647	103,777
User FE	YES	YES	YES	YES
Date FE	YES	YES	YES	YES
Control For Prior	YES	YES	YES	YES

Table A4: Fraction of Positive and Negative Posts and Recent Bitcion Returns

This table presents the association between the fraction of positive and negative posts within a day and Bitcoin returns in the past 14 days. In column (1) and (2), the dependent variable is the fraction of positive posts within one day. In column (3) and (4), the dependent variable is the fraction of negative posts within one day. We calculate the average levels of sentiment in RavenPack news in the past 14 days as controls for news arrivals. We control for market information by controlling for two Bitcoin market variables: Bitcoin Volatility and Number of Transactions. Each of two variable is the cumulative sum in the past 14 days. We multiply each coefficient by 100. The t-statistics (in parentheses) are based on standard errors clustered by date. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
Bitcoin Return in Past 14 Days	2.978***	2.576***	-2.059***	-1.686***
	(3.53)	(3.57)	(-4.00)	(-3.48)
Bitcoin Volatility in Past 14 Days		-13.320***		5.806***
		(-10.03)		(7.19)
Number of Transactions in Past 14 Days		0.197***		-0.057***
		(14.76)		(-8.77)
RavenPack News Sentiment in Past 14 Days		2.665***		-1.706***
		(5.29)		(-5.74)
Adjusted R-Squared	0.005	0.117	0.006	0.043
N	2,280	2,280	2,280	2,280
User FE	NO	NO	NO	NO
Date FE	NO	NO	NO	NO
Controls	YES	YES	YES	YES

Table A5: Real Effect: Shorter Window for Sentiment Contagion

This table presents regression analysis of users' subsequent trading decision on the sentiment change between their Post[0] and Post[1], and we require the gap between Post[0] and Post[1] to be less than 12 hours.In column (1) and (2), the dependent variable is a dummy that equals one if the user's first transaction after publishing post[1] is to buy bitcoins and if this first transaction occurs within a 12-hour window after post [1], and zero otherwise. In column (3) and (4), the dependent variable is a dummy that equals one if the user's first transaction is to buy bitcoins and zero otherwise. Compared to column (1) and (2), in column (3) and (4), we do not restrict a time window for the first transaction after post [1]. The control variables are the same as those in Table 3. In all columns we add user and date fixed effects. We multiply each coefficient by 100. The t-statistics (in parentheses) are calculated based on standard errors clustered by user and date.

	(1) First transaction as buy within 12 hours	(2) First transaction as buy within 12 hours	(3) First transaction as buy	(4) First transaction as buy
Sentiment Change	0.186**	0.245**	0.201	0.426
between Post[0] and Post[1]	(2.27)	(1.98)	(1.22)	(1.50)
RavenPack News Sentiment		-0.417		0.369
between Post[0] and Post[1]		(-1.20)		(0.53)
RavenPack News Sentiment		0.188		-1.892*
24 hours before Post[0]		(0.27)		(-1.80)
RavenPack News Sentiment		-0.738		-0.156
48 hours before Post[0]		(-1.36)		(-0.12)
Bitcoin Return		0.080		6.791
		(0.03)		(1.14)
Bitcoin Volatility		-5.503		-37.162**
		(-0.86)		(-2.39)
Number of Transactions		-0.000		0.000***
		(-0.86)		(2.91)
Forum Sentiment		-0.042		-1.058
		(-0.10)		(-1.55)
Adjusted R-Squared	0.190	0.161	0.435	0.415
N	39,634	23,120	39,634	23,120
Fund FE	YES	YES	YES	YES
Date FE	YES	YES	YES	YES
Control for Prior	YES	YES	YES	YES

Table A6: Post Decisions and Recent Bitcoin Returns

This table presents the association between users' post decisions and Bitcoin returns in the past 14 days. In column (1) and (2), the dependent variable $PostDecisions_{i,t}$ is a dummy variable that equals one if user i makes at least one post on day t and zero otherwise. We focus on users' active days after their registration. In column (3) to (6), we investigate how, conditional on users' post decisions, recent Bitcoin returns affect sentiment in posts . In column (3) and (4), the dependent variable $PositivePosts_{i,t}$ equals to one if user i publishes a positive posts on day t and zero otherwise. In column (5) and (6), the dependent variable $PessimisticPosts_{i,t}$ equals to one if user i publishes a negative posts on day t and zero otherwise. We calculate the average levels of sentiment in RavenPack news in the past 14 days as controls for news arrivals. We control for market information by controlling for two Bitcoin market variables: Bitcoin Volatility and Number of Transactions. Each of two variable is the cumulative sum in the past 14 days. We calculate forum average sentiment as the average sentiment of posts on the Bitcointalk forum published in the past 14 days. In all columns we add date fixed effect. We multiply each coefficient by 100. The t-statistics (in parentheses) are based on standard errors clustered by user and date. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

	(1) Post Decisions	(2) Post Decisions	(3) Positive Posts	(4) Positive Posts	(5) Pessimistic Posts	(6) Pessimistic Posts
Bitcoin Return	0.149***	0.162***	2.158***	0.756***	-1.436***	-0.598***
in Past 28 Days	(6.53)	(6.85)	(7.09)	(2.93)	(-6.23)	(-2.80)
Bitcoin Volatility		0.390***		2.181**		-2.011**
in Past 28 Days		(6.22)		(2.01)		(-2.09)
Number of Transactions		0.002***		-0.032***		0.042***
in Past 28 Days		(3.58)		(-4.95)		(8.65)
RavenPack News Sentiment		-0.043***		2.880***		-2.152***
in Past 28 Days		(-2.77)		(9.27)		(-9.70)
Forum average sentiment		-0.882***		30.038***		-16.831***
in Past 28 Days		(-9.73)		(11.76)		(-8.59)
Adjusted R-Squared	0.073	0.074	0.042	0.043	0.021	0.022
N	60147616	60034776	436,711	436,644	436,711	436,644
User FE	YES	YES	YES	YES	YES	YES
Date FE	NO	NO	NO	NO	NO	NO
Controls	YES	YES	YES	YES	YES	YES