Social Learning and Sentiment Contagion in the Bitcoin Market

Bing Han, Haoyang Liu and Pengfei Sui May 18, 2024

2024 Five-Star Asia Pacific Workshop in Finance

Researchers have long realized the importance of social activities on economic outcomes

• "Man is by nature a social animal" - Aristotle

In his pioneering works, Robert Shiller argues that social learning plays a pivotal role in influencing investors' decisions

- investors update their beliefs through social interactions
- Shiller further proposes a potential link between social learning and asset pricing dynamics

Despite its prominence in the literature, direct empirical evidence on social learning has been limited

• possibly due to data constraints

Study how sentiment spreads over a social network

- Apply **textual analysis** to conversations on Bitcointalk to measure sentiment
- Study: a user's sentiment changes after a conversation relative to before (Senti Change_{i,j,t0→t1})
 - \Leftarrow **Sentiment of other users** in the same conversation

Study how social learning affects trading and market outcomes

- Individual level: sentiment change \Rightarrow direction of trading
- Aggregate level: sentiment contagion \Rightarrow trading volume, volatility and market crash

Why Bitcoin?

- Bitcoin has its origins and development closely tied to social networks.
- The valuation of Bitcoin depends on the aggregate demand from the social network, or **social adoption** (Cong et al. 2020).
- Moreover, the **bubble-like** features of the Bitcoin market are consistent with the long-standing hypothesis on the role of social dynamics in fueling bubbles (Shiller, 1984; Burnside et al., 2016)

Why Bitcointalk?

- Established by Satoshi Nakamoto, the founder of Bitcoin
- Widely covered by Wall Street Journal, Forbes, and Bloomberg

Inefficient social learning:

- Social sentiment $(+) \Rightarrow$ Sentiment change
- Social sentiment $(-)(\mathbf{0}) \Rightarrow$ Future returns

Other evidence for inefficient learning:

- Less sophisticated, less socially connected, and less informed investors are more susceptible
- Intensity of learning is higher on days with higher uncertainty and higher volatility but not on days with more news

Evidence for echo chamber and confirmation bias

Overview of Results: Sentiment contagion affects asset pricing dynamics

- Individual level: sentiment change $(+) \Rightarrow$ direction of trading
- Aggregate level: sentiment contagion $(+) \Rightarrow$ trading volume, volatility and market crash
- During bubble episodes, trading volume is highly correlated with sentiment contagion

- Data
- Evidence for Sentiment Contagion
- Inefficiency of Social Learning
- Social Interactions, Individual Trading and Market Outcomes
- Social Learning and Bubbles

Data

Textual Analysis

We use textual analysis to extract user sentiment revealed in their ${\sf posts}/{\sf messages}$

Our algorithm involves two steps:

- Keyword dictionary: manually label randomly selected 10,000 sentences into 3 categories:
 - positive, neutral and negative. (Baker, Bloom and Davis(2016) and Tetlock (2007))
- Stanford NLP: apply the algorithm to detect
 - the tense of sentences
 - the negative particles (also called negative adverb)

Some labeled examples (out-of-sample accuracy: 85%):

- Bitcoin price will roar (1)
- Bitcoin price won't roar (-1)
- Bitcoin price increased a lot (irrelevant)

Common Words in Our Sample



Bitcoin Transaction Data

• A subsample of investors on Bitcointalk voluntarily published their wallet addresses

Other data sources

- Ravenpack News Analytics: Bitcoin news
- CoinAPI: returns and trading volume at the hourly level
- Google Search Volume Index for Bitcoin

Evidence for Sentiment Contagion

We study consecutive pairs of posts from the same user

• Sentiment Change = ex-post sentiment - prior sentiment

Between a consecutive pair of posts by a user, other users may publish posts in the same thread

• These additional posts are called a **conversation**. The average sentiment is called **social sentiment**

Example of Conversation (Part 1)



Example of Conversation (Part 2)



Study how the average sentiment of the 5 posts in between affects DavidLuziz's sentiment, controlling for news arrivals, Bitcoin market dynamics

Senti Change<sub>*i*,*j*,*t*₀
$$\rightarrow$$
*t*₁ = β_1 Social Sentiment<sub>*i*,*j*,*t*₀ \rightarrow *t*₁
+ γ' Control_{*i*,*t*₀ \rightarrow *t*₁ + Fixed Effects + *u*_{*i*,*t*₁} (1)}</sub></sub>

		Sentiment Change							
	(1)	(2)	(3)	(4)	(5)				
Social Sentiment	8.896*** (21.59)	5.336*** (13.54)	2.844*** (6.83)	3.100*** (6.03)	4.228*** (6.91)				
Prior FE	Х	Х	Х	Х	Х				
User FE		Х	Х	Х					
Day FE			Х	Х					
User X Week FE					Х				
Controls				Х	Х				
Adjusted R-Squared N	0.338 218,268	0.344 212,647	0.344 212,623	0.345 175,750	0.322 143,092				

- A conversation with social sentiment $1\Rightarrow$ Senti Change of 4.228%
- (Full Table)

		Sentiment Change					
	(1)	(2)	(3)	(4)	(5)		
Indirect Social Sentiment	18.783*** (13.67)	7.375*** (5.18)	3.078* (1.87)	3.458* (1.84)	6.740*** (2.86)		
Prior FE	Х	Х	Х	Х	Х		
User FE		Х	Х	Х			
Day FE			Х	Х			
User X Week FE					Х		
Controls	Х	Х	Х	Х	Х		
Adjusted R-Squared N	0.342 97,724	0.348 93,419	0.347 93,164	0.346 75,548	0.327 55,715		

The sentiment of a neighbor's neighbor affects a user's sentiment revision due to the diffusion of sentiment

• sentiment of a neighbor's neighbor is denoted as *Indirect Social* Sentiment

	Sentiment Change							
	(1)	(2)	(3)	(4)	(5)			
Random Conversation	2.682***	1.075***	-0.139	0.398	0.256			
	(10.41)	(4.11)	(-0.51)	(1.20)	(0.64)			
Prior FE	Х	Х	Х	Х	Х			
User FE		Х	Х	Х				
Day FE			Х	Х				
User X Week FE					Х			
Controls				Х	Х			
Adjusted R-Squared N	0.336 257,763	0.344 252,098	0.345 252,077	0.345 198,034	0.321 164,058			

Sentiment changes are NOT affected by the average sentiment in a randomly selected conversation

• Sentiment of a randomly selected conversation is denoted as *Random Conversation*

Inefficiency of Social Learning

Less sophisticated investors are more susceptible to social sentiment

	Naive Users		Centra	Central Users		Less Informed Users	
	(1)	(2)	(3)	(4)	(5)	(6)	
User Feature * Social Sentiment	3.585*** (3.46)	3.321 *** (2.64)	-0.047*** (-3.82)	- 0.047 *** (-3.41)	5.137** (2.24)	7.378 ** (2.44)	
Social Sentiment	1.125 (1.52)	2.477*** (2.89)	4.086*** (7.28)	5.466*** (7.94)	2.304*** (4.10)	3.998*** (5.97)	
User Feature			-0.006 (-0.66)	0.128 (0.91)	-1.576 (-1.28)	-2.069 (-0.76)	
Adjusted R-Squared N	0.345 175,750	0.322 143,092	0.345 175,750	0.322 143,092	0.343 139,873	0.323 118,433	
Prior FE	Х	Х	Х	Х	Х	Х	
User FE	Х		Х		Х		
Day FE	Х		Х		Х		
User X Week FE		Х		Х		Х	
Controls	Х	Х	Х	Х	Х	Х	

• Naive: all users except for those assigned as "legendary" by Bitcointalk

- Central: users participating in more conversations
- Less informed: users whose sentiment less correlated with future returns

17

Sentiment contagion is stronger on more uncertain days, not on days with more news

	Informat	tive Days	High Un	High Uncertainty		oin Volatility
	(1)	(2)	(3)	(4)	(5)	(6)
Episode Feature * Social Sentiment	0.959 (0.92)	0.858 (0.69)	3.714*** (3.01)	3.851 ** (2.36)	2.515*** (2.62)	2.919** (2.50)
Episode Feature		0.448 (0.50)		-1.059 (-1.10)	-0.667 (-0.66)	-1.501* (-1.96)
Social Sentiment	2.664*** (3.76)	3.874*** (4.74)	2.415*** (4.23)	3.613*** (5.42)	1.957*** (2.93)	2.856*** (3.59)
Adjusted R-Squared N	0.345 175,750	0.322 143,092	0.345 175,750	0.322 143,092	0.345 175,750	0.322 143,092
Prior FE	Х	Х	Х	Х	Х	Х
User FE	Х		Х		Х	
Day FE	Х		Х		Х	
User X Week FE		Х		Х		Х
Controls	Х	Х	Х	Х	Х	Х

- Informative: days with more news flow
- · Uncertainty: std of news sentiment on Bitcoin
- Volatility: std of hourly Bitcoin returns

Cumulative Return_{*i*,*j*, $t_1+1 \rightarrow t_1+k$ = β_1 Social Sentiment_{*i*,*j*, $t_0 \rightarrow t_1$}}

+
$$\gamma \text{Control}_{t_0 \rightarrow t_1}$$

$$+ \alpha_i + \gamma_{d(t_0)} + u_{i,t_1}$$

	(1)	(2)	(3)	(4)	(5)
Future returns	6 Hours	24 Hours	48 Hours	72 Hours	168 Hours
Social Sentiment	-0.012	-0.094**	-0.119***	-0.111**	-0.114**
	(-0.38)	(-2.17)	(-2.69)	(-2.31)	(-2.57)
User FE	Х	Х	Х	Х	Х
Date FE	Х	Х	Х	Х	Х
Control For Prior	Х	Х	Х	Х	Х
Adjusted R-Squared	0.268	0.675	0.824	0.878	0.939
Ν	175,750	175,750	175,750	175,750	175748

Social sentiment of most types of investors is misinformed

We decompose social sentiment into two components by each user feature: e.g. social sentiment by informed users versus non-informed users

Cumulative Return_{*i*,*j*, $t_1+1 \rightarrow t_1+k =$}

 β_1 Social Sentiment of a Subsample of Users_{*i*,*j*,*t*₀ \rightarrow *t*₁ + · · ·}

	Inf	ormed	Central		Legendary	
	(1)	(2)	(3)	(4)	(5)	(6)
Social Sentiment	0.007		-0.085**		-0.054	
by Featured Users	(0.08)		(-2.23)		(-1.12)	
Social Sentiment		-0.091**		-0.036		-0.061*
by NonFeatured Users		(-2.27)		(-0.99)		(-1.71)
Adjusted R-Squared	0.678	0.678	0.673	0.695	0.674	0.676
N	61974	112938	159205	119231	138938	161144
User FE	Х	Х	Х	Х	Х	Х
Date FE	Х	Х	Х	Х	Х	Х
Controls	Х	Х	Х	Х	Х	Х

Panel A: Informativeness (future 24-hour returns)

Investors react more to uninformed, central and naive users

Senti Change_{*i*,*j*, $t_0 \rightarrow t_1$ =}

 β_1 Social Sentiment of a Subsample of Users_{*i*,*i*,*t*₀ \rightarrow *t*₁ + · · ·}

Panel B: Response (Sentiment Change)								
	Info	ormed	Cen	tral	Legendary			
	(1)	(2)	(3)	(4)	(5)	(6)		
Social Sentiment	0.483		2.897***		0.797*			
by Featured Users	(0.72)		(5.55)		(1.65)			
Social Sentiment		1.409***		1.925***		2.410***		
by NonFeatured Users		(2.62)		(2.84)		(4.97)		
User FE	Х	Х	Х	Х	Х	Х		
Date FE	Х	Х	Х	Х	Х	Х		
Controls	Х	Х	Х	Х	Х	Х		
Adjusted R-Squared	0.341	0.345	0.345	0.346	0.347	0.345		
N	61,974	112,938	173,477	43,701	138,938	161,144		

Danal D. Daaranaa (Canting ant Channel)

 Unlikely rational given that uninformed, central and naive users' sentiment more negatively predict returns

Echo Chambers (Cookson et al. 2023)

User are more likely to participate in conversations with the **same sentiment as their priors**

 $Pr[Participate Positive_{i,t+k} = 1]_t$

 $= \Phi[\beta_0 + \beta_1 \text{Prior Sentiment}_{i,t} + \gamma_m \text{Control}_{i,t,m} + u_{t+k}], \quad (2)$

	Positive Sentim	ent of First Post	Positive Social Sentiment		
	(1)	(2)	(3)	(4)	
Positive Prior	6.734***	5.721***	4.251***	3.530***	
	(19.64)	(15.97)	(17.84)	(14.37)	
Controls		Х		Х	
Adjusted R-Squared	0.004	0.019	0.002	0.013	
Ν	190,082	131,407	142,871	123,084	

Selective Interpretation of Information

Users react more to conversations with the same sentiment as their priors

• Social sentiment (-) and Social sentiment (+) are the negative and positive components of the social sentiment

	Positiv	e Priors	NonPosit	tive Priors
	(1)	(2)	(3)	(4)
Social Sentiment	4.293***		3.307***	
	(6.08)		(4.70)	
Social Sentiment(+)		5.006***		2.803***
		(4.96)		(2.77)
Social Sentiment(-)		2.742*		4.250***
		(1.65)		(2.75)
User FE	Х	Х	Х	Х
Date FE	Х	Х	Х	Х
Controls	Х	Х	Х	Х
Adjusted R-Squared	0.044	0.044	0.248	0.248
Ν	85,951	85,951	86,230	86,230

Social Interactions, Individual Trading, and Market Outcomes

Individual Trading Data

• Out of the 37,262 users in our sample, 1,284 users voluntarily published their Bitcoin wallet address to protect the security of their Bitcointalk account

We regress trading decisions for individual i on the sentiment change between post[0] and post[1] in conversation j:

$$\begin{array}{lll} \mathsf{Trade}_{i,j,t_0 \to t_1} & = & \beta \mathsf{Senti} \; \mathsf{Change}_{i,j,t_0 \to t_1} + \gamma \mathsf{Control}_{i,j,t_0 \to t_1} \\ & + & \mathsf{Fixed} \; \mathsf{Effects} + u_{t+k} \end{array}$$

Social Learning and Individual Trading

	Unco	nditional	Co	Conditional on trading occurs			
	(1) Probability trade in [t0, t1]	(2) Probability trade in [t1, t1+7days]	(3) Probability buy in [t0, t1]	(4) Probability buy in [t0, t1],t1-t0>2	(5) Probability buy in in [t1, t1+7days]		
Sentiment Change	0.121	-0.039					
	(1.38)	(-0.22)					
Sentiment Change			7.011*	9.244**	0.576**		
			(1.85)	(2.29)	(2.00)		
User FE	Х	Х	Х	Х	Х		
Date FE	Х	Х	Х	Х	Х		
Controls	Х	Х	Х	Х	Х		
Adjusted R-Squared	0.086	0.629	0.252	0.261	0.632		
N	55,297	55,297	203	189	9,284		

- Sentiment change does not unconditionally predict individuals' future trading decisions
- Conditional on trading, sentiment change predicts the trade direction
- Consistent with Giglio et al. 2021

Aggregate sentiment contagion predicts trading volume and volatility

We construct a daily Sentiment Contagion Intensity (SCI) index to capture the aggregate intensity of sentiment contagion in social interactions. (Details of SCI)

$$\mathsf{Market Outcome}_{t+N} = \beta_0 + \beta_1 SCI_{i,t} + \Sigma_m \gamma_m \mathsf{Controls}_{t,m} + u_{t+k} \quad (3)$$

	Abnormal Tr	ading Volume	Return '	Volatility
	(1)	(2)	(3)	(4)
	1 Day	7 Days	1 Day	7 Days
SCI	401.246***	304.257***	0.536***	0.385***
	(3.26)	(2.69)	(5.55)	(4.62)
Controls	X	X	X	X
Adjusted R-Squared	0.025	0.079	0.262	0.385
N	3,664	3,663	3,664	3,663

Aggregate sentiment contagion predicts market crash

Hirshleifer (2020): the error-prone feature regarding the impact of social interactions on the market

• SCI index also predicts market crashes

 $\mathsf{Prob}(\mathsf{Market } \mathsf{Crash}_{t+N}) = \beta_0 + \beta_1 SCI_{i,t} + \Sigma_m \gamma_m \mathsf{Controls}_{t,m} + u_{t+k}$ (4)

	Crash (below 1% perc.)		Crash (belo	ow 5% perc.)
	(1) (2)		(3)	(4)
	1 day	7 days	1 day	7 days
SCI	0.451** (2.22)	1.824*** (4.55)	2.029*** (4.58)	8.361*** (10.59)
Controls	Х	Х	Х	Х
Adjusted R-Squared	0.125	0.119	0.052	0.068
Ν	3664	3664	3664	3664

Social Learning and Bubbles

Identify Bitcoin Bubbles

We identify bubbles by conforming to Fama's notion that a bubble, if it does exist, is associated with a substantial price run-up

- we select days with a cumulative Bitcoin return higher than 200% over the past quarter
- we select the peak day and the day two quarters prior to the peak day as the start date

In total, we identified four bubble episodes:

- June 4, 2013 to December 4, 2013
- June 16, 2017 to December 16, 2017
- December 26, 2018 to June 26, 2019.
- September 13, 2020 to March 13, 2021

More Results on Bubbles

Trading Volume in Bubbles

Correlation between the number of positively infected users and future trading volume in Bubble Episode 1 (June 2013 to December 2013):

• 0.669 (p-value = 0.00)



Relying on evidence from an online investor community, this paper presents direct evidence for the role of social learning

- Social learning leads to sentiment contagion in social interactions
- Social learning process is inefficient, as investors respond positively to social sentiment, but social sentiment does not positively predict returns
- Social learning influences individuals' trading and market outcomes
- Social learning is connected to bubbles

Such novel data may help us understand other important questions in social finance

Appendix

Full Table

	Sentiment Change				
	(1)	(2)	(3)	(4)	(5)
Social Sentiment	8.896*** (21.59)	5.336*** (13.54)	2.844*** (6.83)	3.100*** (6.03)	4.228*** (6.91)
RavenPack News Sentiment between Post[0] and Post[1]				-0.037 (-0.04)	1.569* (1.70)
RavenPack News Sentiment 24 hours before Post[0]				1.378 (1.44)	1.681* (1.86)
RavenPack News Sentiment 48 hours before Post[0]				-1.021 (-1.02)	-0.007 (-0.01)
Bitcoin Return				30.727*** (5.90)	32.487** (5.62)
Bitcoin Volatility				13.020 (0.81)	5.273 (0.30)
Forum Sentiment				-4.691*** (-4.43)	1.623 (1.35)
Prior FE	YES	YES	YES	YES	YES
User FE	NO	YES	YES	YES	NO
Day FE	NO	NO	YES	YES	NO
User X Week FE	NO	NO	NO	NO	YES
Controls	NO	NO	NO	YES	YES
Adjusted R-Squared	0.338	0.344	0.344	0.345	0.322
N	218,268	212,647	212,623	175,750	143,092

Back

Construction of SCI

We construct a daily Sentiment Contagion Intensity (SCI) index to capture the aggregate intensity of sentiment contagion in social interactions

SCI is generated in two steps:

- Within each day, we count the number of investors who change their sentiment in the same direction as the social sentiment
- We remove the time trend and seasonality by regressing the series for affected users on the weekday and year-month-pair indicators

	Bubble Episode		Non-Bubble Episode		Bubble Episode minus Non-Bubble Episode	
Features	Mean	Std	Mean	Std	Difference	t-statistic
Panel A: Market Variables						
Daily Return	4.050	17.66	0.491	15.72	3.56	5.352
Return Volatility	0.040	0.03	0.037	0.04	0.003	2.413
Dollar Volume	94.071	148.80	47.503	77.63	46.568	11.784
News Sentiment	0.168	0.42	0.040	0.43	0.127	5.894
Google Search	0.085	0.48	0.017	0.40	0.068	3.961

In bubble episodes:

- daily return is more than 8 times higher
- return volatility within a day increases by 8.108%
- dollar volume nearly doubles
- news sentiment is more than four times higher
- Google Search volume surges

Social Interactions in Bubbles

			Bubble Episode			Episode
	Bubble Episode		Non-Bubble Episode		minus Non-Bubble Episode	
Features	Mean	Std	Mean	Std	Difference	t-statistic
Panel B: Social Interactions						
Average Sentiment	0.305	0.11	0.272	0.12	0.033	6.761
Std of Sentiment	0.657	0.05	0.672	0.05	-0.015	-6.635
Number of Posts	188.764	147.95	175.713	163.66	13.052	1.968
Number of Users	133.504	105.61	116.261	114.97	17.243	3.692
Number of Positive Posts	104.671	87.03	90.834	87.47	13.837	3.837
Fraction of Positive Posts	0.543	0.09	0.516	0.09	0.028	7.563
Fraction of Sophisticated Users	0.301	0.10	0.359	0.14	-0.058	-10.523

In bubbles:

- user sentiment becomes more optimistic and less dispersed
- · intensity in social interactions becomes stronger
- the number of posts with positive sentiment increases disproportionately
- a disproportionately large number of novice investors participate in social interactions

Elevated Contagion of Optimism

					Bubble Episode		
	Bubble Episode		Non-Bubble Episode		minus Non-Bubble Episode		
Features	Mean	Std	Mean	Std	Difference	t-statistic	
Panel C: Sentiment Contagion							
Number of Positively Infected Users	16.704	16.24	15.335	17.33	1.369	1.938	
Number of Negatively Infected Users	3.203	4.63	3.813	5.84	-0.61	-2.629	

In bubbles:

- the number of positively infected users significantly increases by 8.927%, while the number of negatively infected users on average experiences a significant decrease of 15.998%.
- Why?
 - users would frequently run into conversations with an overall positive sentiment
 - there is an increasing participation of naive investors in discussions
 - sentiment contagion becomes stronger during periods of high uncertainty

Back

Trading Volume in Bubbles

Correlation between the number of positively infected users and future trading volume in Bubble Episode 4 (September 2020 to March 2021):

• 0.335 (p-value = 0.00)



	(1)	(2)	(3)	(4)	(5)	(6)
	6 Hours	6 Hours	24 Hours	24 Hours	72 Hours	72 Hours
Optimism in Summation	0.179* (1.95)		0.419** (2.16)		1.312*** (3.58)	
Optimism in Fraction		0.414** (2.24)		1.182*** (3.24)		2.782 ^{***} (4.38)
RavenPack News Sentiment	0.039	-0.184	0.519	-0.215	1.002	-0.751
in Past 7*24 Hours	(0.06)	(-0.29)	(0.44)	(-0.19)	(0.50)	(-0.39)
Bitcoin Return Volatility	170.494***	167.886***	441.053***	422.925***	1021.164***	981.951***
in Past 7*24 Hours	(4.78)	(4.76)	(6.16)	(6.27)	(8.16)	(8.18)
Cumulative Bitcoin Return	-4.347	-3.308	-12.444*	-9.780	-18.378	-10.586
in Past 7*24 Hours	(-1.02)	(-0.78)	(-1.65)	(-1.27)	(-1.44)	(-0.82)
Adjusted R-Squared	0.109	0.116	0.127	0.141	0.140	0.153
N	17953	17953	17953	17953	17953	17953