Trading the Dream: Does Social Media Affect Investors' Activity - The Story of Twitter, Telegram and Reddit

Tiran Rothman[1]

[1] Wizo Haifa Academic College, Haifa, Israel

Correspondence: Tiran Rothman, Wizo Haifa Academic College, Haifa, Israel.

Abstract

Bitcoin has received much attention in the media and by investors in recent years, although there remains scepticism and a lack of understanding of this cryptocurrency. I add to the literature on Bitcoin by studying How and if social media platforms such as Twitter, Telegram and Reddit are affecting Bitcoin's prices and volume. Findings confirm that volumes of exchanged tweets may predict the fluctuations of Bitcoin's prices and mainly Bitcoin volume. I also find that Reddit and Telegram have greater impact on Bitcoin volume rather than Twitter. Results indicate that information about the use of Twitter and other social media platforms can assist in tracking real world behavior and may even predict real financial markets trend.

Keywords: Bitcoin, behavioural finance, cryptocurrency

JEL Classification: G14, G12

1. Introduction

Bitcoin symbolizes a radical change in financial systems, attracting a large number of users and a lot of media attention. In this work, I focus on Bitcoin and investigate if the Bitcoin's price spread is related to social media volumes; and explore how positive or negative social media sentiments affect in Bitcoin's investors' activity.

Social media can be a powerful tool in predicting investors` activity. It has been commercially exploited for goals such as automatically extracting customer opinions about products or brands, to find which aspects are liked and which are disliked Thelwall (2011). Ye and Wu (2010) demonstrate how particularly interesting is the influence of Twitter users and the propagation of the information related to their tweets. Twitter had become the world's seventh most popular website by March 2015 (Alexa, 2014). Twitter is an online social networking website and microblogging service that allows users to post and read text-based messages of up to 140 characters, known as "tweets". Launched in July of 2006 by Jack Dorsey, Twitter is now in the top 10 most visited internet sites with a total amount of 645,750,000 registered users. The short format of a tweet is a defined characteristic of the service, allowing informal collaboration and quick information sharing. For business, Twitter can be used to broadcast company's latest news, posts, read comments of the customers or interact with them. A communicative feature of Twitter is the hashtag: a metatag beginning with the character #, designed to help others find a post. Twitter is a rich source of real-time information regarding current societal trends and opinions.

There are also studies that report another use of Twitter, namely as a possible predictor of market trends. Indeed, in 2010, a publication of the professor Johan Bollen (2011) showed that combining information on Wall Street with the millions of Tweets and posts, makes possible to anticipate financial performance. In this work, Granger causality analysis and a Self-Organizing Fuzzy Neural Network are used to investigate the hypothesis that public mood states, as measured by the OpinionFinder and GPOMS mood time series, are predictive of changes in DJIA closing values. The analysis of Tweets made by Bollen would have had 87% of chance to successfully predict prices of the stock, 3 or 4 days in advance. This study and analysis of millions of posts on Twitter represents a thermometer of emotions, on a large scale, which reflects the whole of society. Earlier studies had found that blogs can be used to evaluate public mood, and that tweets about movies can predict box office sales.

Investigating the literature related to different uses of social media, and Twitter in particular, we collected information about the use of Twitter for seeking real world emotions that could predict real financial markets trend. In their paper, Rao and Srivastava investigate the complex relationship between tweet board literature (like

bullishness, volume, agreement etc) with the financial market instruments (like volatility, trading volume and stock price). The Bitcoin represents an important new phenomenon in financial markets. Garcia et al. (2014) show the interdependence between social signals and price in the Bitcoin economy, namely a social feedback cycle based on word-of-mouth effect and a user-driven adoption cycle. They provide evidence that Bitcoin's growing popularity causes increasing search volumes, which in turn result a higher social media activity about Bitcoin.

Mai et al. (2018) examine predictive relationships between social media and Bitcoin returns by considering the relative effect of different social media platforms (Internet forum vs. microblogging) and the dynamics of the resulting relationships using vector autoregressive and vector error correction models. More interest inspire the purchase of bitcoins by users, driving the prices up, which eventually feeds back on the search volumes.

2. Literature Overview

The literature on Bitcoin was initially dominated by studies on the safety, ethical and legal aspects of Bitcoin. Recently, some literature has examined Bitcoin from an economic viewpoint. Selgin (2015) argued that investors have employed Bitcoin as currency as well as for investment purposes, although, they claimed that Bitcoin should be seen as a speculative commodity rather than a currency. Dwyer (2015) finds that the average monthly volatility of Bitcoin is higher than that for gold or a set of foreign currencies, and the lowest monthly volatilities for Bitcoin are less than the highest monthly volatility for gold and currencies.

Cheah and Fry (2015) argue that if Bitcoin were a true unit or account, or a form of store of value, it would not display such volatility expressed by bubbles and crashes. Cheung et al (2015) show the existence of bubbles in the bitcoin market over the period and find a number of short-lived bubbles but also three huge bubbles, the last of which led to the demise of the Mt Gox exchange. Brière et al (2015) show that Bitcoin offers significant diversification benefits for investors while Dyhrberg (2016a; 2016b) show that Bitcoin has similar hedging capabilities as gold and the dollar, and as such can be employed for risk management.

Fry and Cheah (2016) develop a model to reveal that Bitcoin and Ripple are characterized by negative bubbles. Bouri et al. (2017) scrutinize hedge and safe haven properties of Bitcoin vis-`a-vis several stock, bonds and currency indices around the world. Its main finding is that the cryptocurrency is only useful as a diversifier device, but not as a hedge instrument. Finally, Balcilar et al. (2017) detect nonlinearities in the return-volume relationship, which allfows for return prediction. None of these studies analyzed, as to our knowledge, the correlation among this 'basket of coins'.

I focus on Twitter as a leading social media platform, rich source of real-time information regarding current social trends and opinions, however, other social media platforms have started to emerge as a replacement to Twitter. While there are several studies that report use of Twitter, namely as a possible predictor of market trends, as far as I know, none has explored other social media platforms to cryptocurrency market activity. Bollen (2010) showed that combining information on Wall Street with the millions of Tweets and posts makes possible to anticipate financial performance. The analysis of Tweets made by Bollen would have had 87% of chance to successfully predict prices of the stock, 3 or 4 days in advance. Rao and Srivastava (2012) investigate the complex relationship between tweet board literature (like bullishness, volume, agreement etc) with the financial market instruments (like volatility, trading volume and stock price). Mai and Hranac (2013) examine predictive relationships between social media and Bitcoin returns by considering the relative effect of different social media platforms (Internet forum vs. microblogging) and the dynamics of the resulting relationships using vector autoregressive and vector error correction models.

Rothman (2018) explored the digital coins eco-system correlations based on an hourly time interval. The findings show that bitcoin price and volume is not correlated with most of the traded digital coins while several digital coins are highly and significantly correlated with other coins. This has implications for risk management and financial engineering (such as bitcoin derivatives)-both from an investor's as well as from a regulator's point of view.

3. Methodology

3.1 Sentiment Analysis

In recent years, there is a wide collection of research surrounding machine learning techniques, in order to extract and identify subjective information in texts. This area is known as sentiment analysis or opinion mining. Sentiment techniques are able to extract indicators of public mood directly from social media content. Similar to Go et al. (2009) that affirmed the strength of the sentiment analysis applied to the Twitter domain by using similar machine learning techniques to classifying the sentiment of tweets, I chose to use automated sentiment analysis techniques to identify the sentiments of tweets in the matter of Bitcoin.

Since the goal of this research is neither to develop a new sentiment analysis nor to improve an existing one, I use "SentiStrenght", a tool developed by a team of researchers in the UK that demonstrated good outcomes (see Kim, 2009, Thelwall et al., 2013, Thelwall 2017) (Note 1). SentiStrength estimates the strength of positive and negative sentiments in short texts. It is based on a dictionary of sentiment words, each one associated with a weight, which is its sentiment strength.

3.2 Data Collection

I collect tweets from November 2017 till August 2018 on an hourly basis. Tweets are available and are easily retrieved making use of Twitter Application Programming Interface (API) (Note 2). Composing the hashtag #Bitcoin or @bitcoin, I'm able to gather all tweets that mentioned the analyzed subject and create a fundamental database of intra-day prices and volumes. I then run SentiStrenght tool and analyze a collection of tweets regarding Bitcoin, posted on Twitter between November 2017 and August 2018. During this time 1,924,891 tweets were collected.

The tweets were analyzed to determine its identifier, the date-time of the submission, its type, and its text content, which is limited to 140 characters. I then use SentiStrenght to evaluate comments extracted from Twitter. Given as input all tweets, the system assigned a score for each comment: 1 if the comment is positive; -1 if the comment is negative; 0 if the comment is neutral. I also pull volume data for the exact time from Reddit and Telegram, i.e. how many social comments were in these social media platforms. Social volume is a new tool created by Solume, Fintech start-up, who measures social volume where Twitter, Telegram and Reddit have equal balance in volume (Note 3).

To determent whether variation in social media activities have an effect on Bitcoin price and volume, I estimate the following models:

$$\Delta Y_{it} = b_0 + b_1 \Delta negative_{it} + b_2 \Delta positive_{it} + b_3 \Delta reddit_{it} + b_4 \Delta sentiment_{it} + b_5 \Delta telegram_{it}$$
$$+ b_5 \Delta twitter_{it} + b_6 \Delta social_{volume_{it}} + \varepsilon_{it} \tag{1}$$

where Y_{it} is the outcome variable (price and volume) of digital coin ID i at time t. the variables of social_volume$_{it}$, negative$_{it}$, positive$_{it}$ and sentiment$_{it}$ represent the overall social volume, the negative social volume, the positive social volume and the sentiment social volume of each coin i at time t. reddit$_{it}$, telegram$_{it}$ and twitter$_{it}$ corresponds with the social volume of the social network.

Main goal is to study the impact of the changes in social volumes on the changes in Bitcoin prices or volume, therefore each variable is defined as the difference as follows:

$$\Delta x = X_t / X_{t-1} - 1$$

The depended variable of volume is defined as the difference above, yet the depended variable of price is defined as the difference of the log price:

$$\Delta volume_{it} = volume_t / volume_{t-1} - 1$$
$$\Delta price_{it} = log (price_t / price_{t-1})$$

I also investigate casualty between changes in social volume and in price and volume to determine rather social media influenced market behavior or the opposite. To wit I conduct Granger causality test (Granger, 1980) for a time difference of no lag (model 1), 1 hour (model 2), 1 day (model 3), 3 days (model 4) and 1 week (model 5). All Bitcoin hourly data is extracted from www.Binace.com.

4. Results

The results confirm that volumes of exchanged tweets may predict the fluctuations of Bitcoin's price. Furthermore, comparison between tweets with a positive/negative mood and trend of Bitcoin's price seems to prove this behavior.

Prices - Table A below indicates the impact of social media on Bitcoin's price. The coefficient for social volume on prices is positive and significant in some cases. Negative sentiment in Twitter affects Bitcoin prices (Table A, model 1), however, all other social media do not have a significant affect. Observing casualty, I find that mainly Telegram has vast and significant effect on Bitcoin prices (Table A, model 4), while all other social media platform do not seems to have any robust effect on prices.

Volume - Table B shows also results of casualty test on volume (models 2 to 5). Social media seems to have more effect on Bitcoin volume than on its price. Coefficients for social volume is positive in models 1,2,4 and 5 yet it is

only significant in models 1 and 2, social media affect in the very short time Bitcoin's volume. However, observing Reddit and Telegram, I find robust and significant effect on Bitcoin's volume unlike Twitter effect.

Table A. Prices

	5	4	3	2	1
negative_d	-0.0000303	-0.000178	0.0000561	-0.0000142	**-0.000943**[**]
	(-0.08)	(-0.48)	(0.15)	(-0.04)	**(-2.63)**
positive_d	0.000251	-0.000199	-0.000175	0.000339	**0.000974**
	(0.35)	(-0.28)	(-0.25)	(0.48)	**(1.39)**
reddit_d	**-0.00128**	0.000602	0.000821	0.000118	**0.000938**
	(-1.17)	(0.55)	(0.76)	(0.11)	**(0.87)**
telegram_d	-0.00000877	**0.000861**[*]	0.000189	-0.000145	**0.000313**
	(-0.03)	**(2.54)**	(0.56)	(-0.43)	**(0.94)**
twitter_d	0.00136	0.00156	0.000257	-0.000245	**0.00151**
	(0.91)	(1.06)	(0.18)	(-0.17)	**(1.04)**
sentiment_d	-0.0000486	-0.0000179	-0.0000103	-0.0000296	-0.0000320
social_volume_d	0.000357	**-0.00268**	-0.000456	-0.000105	**-0.00233**
	(0.17)	**(-1.30)**	(-0.22)	(-0.05)	**(-1.15)**
_cons	-0.000144	-0.000148	-0.000118	-0.0000718	-0.0000674
	(-0.80)	(-0.82)	(-0.65)	(-0.40)	(-0.38)
N	5793	5889	5937	5960	5961

t statistics in parentheses

[*] $p < 0.05$, [**] $p < 0.01$, [***] $p < 0.001$

Table A shows 5 different models of the relation between Bitcoin hourly price and social media platform. Models: no lag (model 1), 1 hour (model 2), 1 day (model 3), 3 days (model 4) and 1 week (model *5)*

Table B. Volume

	5	4	3	2	1
negative_d	**0.0206**	-0.0147	**0.0461**[*]	-0.00792	-0.115
	(1.06)	(-0.78)	**(2.41)**	(-0.42)	(-0.12)
positive_d	0.0193	**0.0779**[*]	**0.0370**	0.0140	-0.915
	(0.52)	**(2.17)**	**(1.01)**	(0.38)	(-0.48)
reddit_d	0.00621	**0.0682**	0.00213	**0.0805**	**8.231**[**]
	(0.11)	**(1.24)**	(0.04)	**(1.43)**	**(2.77)**
telegram_d	**0.0172**	0.0262	0.0133	**-0.0443**[*]	-0.452
	(1.00)	**(1.53)**	(0.75)	**(-2.54)**	(-0.49)
twitter_d	-0.0523	-0.0654	-0.0679	**0.0847**	**10.84**[**]
	(-0.68)	(-0.88)	(-0.89)	**(1.12)**	**(2.72)**
sentiment_d	**-0.00643**	-0.00117	**0.00645**	0.00412	-0.0940

	(-1.46)	(-0.26)	**(1.43)**	(0.91)	(-0.39)
social_volume_d	0.0789	0.0819	0.0627	**-0.231***	**-9.732**
	(0.74)	(0.79)	(0.59)	**(-2.18)**	**(-1.74)**
_cons	0.116***	0.114***	0.113***	0.122***	0.879
	(12.53)	(12.51)	(12.03)	(13.08)	(1.78)
N	5742	5838	5886	5909	5961

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B shows 5 different models of the relation between Bitcoin hourly volume and social media platform. Models: no lag (model 1), 1 hour (model 2), 1 day (model 3), 3 days (model 4) and 1 week (model 5)

5. Discussion

This study and analysis of millions of posts on Twitter, Telegram and Reddit represents a thermometer of investors' behavior, on a large scale, similar to earlier studies that had found that blogs can be used to evaluate public mood. Results indicates that information about the use of Twitter and other social media platforms can assist in tracking real world behavior and may even predict real financial markets trend. I find that social media can mainly affect Bitcoin volume rather than prices and that less known 'new' social media platforms, such as Reddit and Telegram, have greater impact on investors' activity than Twitter has.

To conclude, this research contributes to the growing literature on cryptocurrency and investors' activity around them. This 'new world' of innovative social media platforms may have crucial part in future trading platforms using also social media as a thermometer. Future research may be in investigating investors' activity with other cryptocurrencies.

References

Alexa. (n.d.). Top sites: The top 500 sites on the Web. Retrieved September 30, 2014, from http://www.alexa.com/topsites/global

Balcilar, M., Bouri, E., Gupta, R., & Roubaud, D. (2017). Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling, 64*, 74-81. https://doi.org/10.1016/j.econmod.2017.03.019

Bollen, J., Mao, H., & Pepe, A. (2011). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. *Icwsm, 11*, 450-453.

Briere, M., Oosterlinck, K., & Szafarz, A. (2015). Virtual currency, tangible return: Portfolio diversification with bitcoin. *Journal of Asset Management, 16*(6), 365-373. https://doi.org/10.1057/jam.2015.5

Cheah, E. T., & Fry, J. (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters, 130,* 32-36. https://doi.org/10.1016/j.econlet.2015.02.029

Dyhrberg, A. H. (2016). Bitcoin, gold and the dollar-A GARCH volatility analysis. *Finance Research Letters, 16*, 85-92. https://doi.org/10.1016/j.frl.2015.10.008

Dyhrberg, A. H. (2016). Hedging capabilities of bitcoin. Is it the virtual gold?. *Finance Research Letters, 16,* 139-144. https://doi.org/10.1016/j.frl.2015.10.025

Garcia, D., Tessone, C. J., Mavrodiev, P., & Perony, N. (2014). The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy. *Journal of the Royal Society Interface, 11*(99), 20140623. https://doi.org/10.1098/rsif.2014.0623

Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N Project Report, 1*(12), Stanford.

Granger, C. W. (1980). Testing for causality: a personal viewpoint. *Journal of Economic Dynamics and Control, 2*, 329-352. https://doi.org/10.1016/0165-1889(80)90069-X

Kim, J. H. (2009b). Automatic variance ratio test under conditional heteroscedasticity. *Finance Research Letters, 3*, 179-185. https://doi.org/10.1016/j.frl.2009.04.003

Mai, E., & Hranac, R. (2013). Twitter interactions as a data source for transportation incidents. *Proc. Transportation Research Board 92nd Ann. Meeting (No. 13-1636)*. https://doi.org/10.1080/07421222.2018.1440774

Mai, F., Shan, Z., Bai, Q., Wang, X., & Chiang, R. H. (2018). How does social media impact Bitcoin value? A test of the silent majority hypothesis. *Journal of Management Information Systems, 35*(1), 19-52.

Rao, T., & Srivastava, S. (2012, August). Analyzing stock market movements using twitter sentiment analysis. *Proceedings of the 2012 international conference on advances in social networks analysis and mining (ASONAM 2012)*, pp. 119-123. IEEE Computer Society.

Rothman, T. (2018). Digital Coins Eco-System: Analyzing the Network of Prices and Volume. *Journal of Finance and Economics, 6*(6), 250-253. https://doi.org/10.12691/jfe-6-6-7

Selgin, G. (2015). Synthetic commodity money. *Journal of Financial Stability, 17*, 92-99. https://doi.org/10.1016/j.jfs.2014.07.002

Thelwall, M. (2017). The Heart and soul of the web? Sentiment strength detection in the social web with SentiStrength. *Cyberemotions*, pp. 119-134. Springer, Cham. https://doi.org/10.1007/978-3-319-43639-5_7

Thelwall, M., & Buckley, K. (2013). Topic-based sentiment analysis for the social web: The role of mood and issue-related words. *Journal of the American Society for Information Science and Technology, 64*(8), 1608-1617. https://doi.org/10.1002/asi.22872

Ye, S., & Wu, F. (2010, August). Estimating the size of online social networks. *Proc. of the IEEE 2nd Intl. Conf. on Social Computing*, pp.169-176. https://doi.org/10.1109/SocialCom.2010.32

Notes

Note 1. http://sentistrength.wlv.ac.uk/

Note 2. https://developer.twitter.com/en/docs/api-reference-index.html

Note 3. https://solume.io/