What makes cryptocurrencies special? Investor sentiment and return predictability during the bubble *

Cathy Yi-Hsuan Chen †, Roméo Després ‡, Li Guo §, and Thomas Renault ¶

July 10, 2019

Abstract

The 2017 bubble on the cryptocurrency market recalls our memory in the dot-com bubble, during which hard-to-measure fundamentals and investors' illusion for brand new technologies led to overvalued prices. Benefiting from the massive increase in the volume of messages published on social media and message boards, we examine the impact of investor sentiment, conditional on bubble regimes, on cryptocurrencies aggregate return prediction. Constructing a crypto-specific lexicon and using a local-momentum autoregression model, we find that the sentiment effect is prolonged and sustained during the bubble while it turns out a reversal effect once the bubble collapsed. The out-of-sample analysis along with portfolio analysis is conducted in this study. When measuring investor sentiment for a new type of asset such as cryptocurrencies, we highlight that the impact of investor sentiment on cryptocurrency returns is conditional on bubble regimes.

Keywords: Cryptocurrency; Sentiment; Bubble; Return Predictability

JEL Classification: G02; G10; G12

*We thank to Tim Loughran for his valuable comments.

†Adam Smith Business School, University of Glasgow, Glasgow, G12 8QQ, UK (*e-mail: CathyYi-Hsuan.Chen@glasgow.ac.uk*)

‡École Normale Supérieure de Rennes, Campus de Ker Lann, Avenue Robert Schuman, 35170 Bruz, France (*e-mail: romeo.despres@ens-rennes.fr*)

§Lee Kong Chian School of Business, Singapore Management University, 50 Stamford Road, Singapore 178899 (*e-mail: liguo.2014@pbs.smu.edu.sg*)

¶Paris 1 Panthéon-Sorbonne University, 17 rue de la Sorbonne, 75005 Paris, France (*e-mail: thomas.renault@univ-paris1.fr*)

1 Introduction

Following the very large price movements in 2017 and 2018, academics and practitioners have started paying strong attention to cryptocurrencies such as Bitcoin and Ethereum. Recent research provides empirical evidence showing that traditional asset pricing models and standard risk factors do not help explaining cryptocurrencies returns (Liu and Tsyvinski, 2018). As there is no fundamental information, such as earnings, dividend and other types of cash flows, the price discovery process of cryptocurrencies is extremely complex. According to Cheah and Fry (2015), the fundamental value of Bitcoin is equal to zero, in such a way that all financial models relying on the well-characterized fundamental values as a reference point cannot be applied to cryptocurrencies. Theoretically, if the market is driven by individual investors who possess higher risk preference, and in the presence of limits to arbitrage (e.g. a short-sale constraint), sentiment-driven noise traders should play an important role in the price discovery process (De Long et al., 1990a). While media often reports investor irrationality, market regulation, and market manipulation as the main drivers of the cryptocurrency market, what ultimately causes all the fluctuations in the price of this new class of assets remains an open question.

In this paper, we examine the relation between individual investor sentiment on social media and cryptocurrencies return. On traditional markets, such as the stock market, investor overoptimism can generate short-term temporary price deviation from the fundamental value (Tetlock, 2007; Tetlock et al., 2008), particularly during economic recessions (Garcia, 2013). Thus, an increase in investor optimism is normally followed by a price reversal, as identified, among others, by Sun et al. (2016) and Renault (2017). However, previous results based on the analysis of traditional markets are not generalizable to the cryptomarket, mainly concerning the permanent versus the temporary effect of investor sentiment on cryptocurrency prices. On one hand, higher investor sentiment can lead to an increase in the adoption rate of cryptocurrencies by increasing confidence and awareness. As confidence is one of the bases of fiat money, and if we consider cryptocurrency as a new type of money, investor sentiment (confidence) can sustain permanently the price of cryptocurrencies. On the other hand, higher investor sentiment can also be a sign, as in the stock market, of an irrational increase in the demand for the speculative asset, and, in that case, should be followed by a price reversal.

Apart from the aforementioned permanent and temporary effect, a prolonged sentiment effect may emerge especially during the bubble period. In such a period, the tendency reverting to the fundamental value is restricted, prolonging the sentiment effect until bubble bursts. As during the observed dot-com bubble in 2000, excessive speculations and illusions about prospects might have led to overvalued prices on the cryptocurrency market. This phenomenon (exuberant sentiment along with over-valuation on asset prices) lasts for a while and has been characterized as a speculative bubble. The revolutionary technologies being mentioned in the dot-com period and introduced to cryptocurrency make the estimation of fundamental value challenging given investors' limited sophistication. Similar to the increase in stock price following the announcement of corporate name changes to Internet-related dot-com names (see Cooper et al. (2001)), investors' irrational fantasy also led to market anomalies during the cryptocurrency bubble, such as when the stock price of a company named "Long Island Iced Tea" increased by 289% in a day when the company changed its name to "Long Blockchain" in December 2017, at the pic of the bubble.[1]

In this paper, we differ from the literature exploring the equilibrium price of cryptocurrency based on a proxy of hard information (Jermann, 2018; Athey et al., 2016; Pagnotta and Buraschi, 2018; Detzel et al., 2018) and from the literature on the economic role of cryptocurrency (Böhme et al., 2015; Harvey, 2014; Balvers and McDonald, 2017; Easley et al., 2017; Yermack, 2017) by emphasizing the importance of soft information and the need to adopt a specific approach to measure sentiment about cryptocurrencies.[2] We contribute to the literature by creating a crypto-specific lexicon to enhance the accuracy of sentiment quantification tailored for crypto assets. Many domain-specific terms, such as blockchain, ICO, hackers, wallet, shitcoin, binance, and "hodl", are not covered in existing financial or psychological dictionaries. Online investors also use new "emojis" such as 🚀 (positive) and 💩 (negative) when talking about cryptocurrencies, which are also not collected in traditional dictionaries. Furthermore, the messages published by online investors on social media are usually shorter and less formal than the content published on traditional media, making the correct classification of tone difficult (Loughran and McDonald, 2016). In that regard, developing a new lexicon is of utmost importance to properly capture the sentiment relative to the cryptocurrency market approaches.

[1] "Long Island Iced Tea Soars After Changing Its Name to Long Blockchain", Bloomberg, 21 December 2017

[2] Textual analysis has been widely used for market participants to collect, process, and transmit soft information (Loughran and McDonald, 2016); Soft information provides incremental explanatory power on firms' future performance, especially when hard information or fundamental information is incomplete or biased (see Tetlock et al. (2008); Lerman and Livnat (2010); Feldman et al. (2010); Loughran and McDonald (2011)).

To construct our lexicon, we follow the methodology of Oliveira et al. (2016) by analyzing a novel dataset of more than 1 million messages related to 425 cryptocurrencies posted on the microblogging platform Stocktwits during 5 years (Jan. 2014 - Dec. 2018). We compare our results to a classification based on the Loughran and McDonald (2011) (LM) lexicon. We find that, in comparison with the LM lexicon, the crypto-specific lexicon achieves 32% higher accuracy in terms of out-of-sample classification, confirming the necessity of using a specific lexicon to measure sentiment on a specific market.

Tremendous efforts on data collection have been made in this study. We use two very large datasets of millions of messages published on two different websites: one social media (StockTwits) and one message boards (Reddit). We choose to analyze two sources of data as StockTwits is mostly related to the speculative part of the cryptocurrency market (crypto as an asset), while the messages on Reddit are much more diverse, and can cover topics related to the technology (blockchain) or the economic role of cryptocurrencies as money. We then construct daily sentiment indicators for each source of data and these indicators afterward are employed in the empirical analysis. Contrary to the vast majority of papers on the literature who focus on a subset of few cryptocurrencies (Bitcoin, Ethereum, Ripple, Monero, Litecoin...), we choose to analyze the aggregate cryptocurrency market by considering the CRyptocurrency IndeX (CRIX) as the weighted return of a large basket of cryptocurrency, constructed by Trimborn and Härdle (2018). Our findings can, therefore, be generalized more easily to the cryptocurrency market as a whole.

Then, we examine the tendency of prices to converge to their fundamental value in the long-run *before* and *post*-bubble period, respectively. We employ the local-momentum autoregression model proposed by Duan (2016) to characterize a parsimonious autoregressive model in the time series of cryptocurrency market return that is globally mean-reverting into its fundamental process but locally driven by momentum. During the bubble, strong local momentum in price dynamics can be attributed to positive feedback trading by noise traders' sentiment (De Long et al., 1990b). Given this fact, a deviation from the fundamental may take longer. We show, through a local-momentum autoregression model, a limited tendency to the latent central factor before and during the bubble, but a recovered tendency after the bubble. To zoom in the short-run momentum phenomena, we conduct a vector autoregression analysis between the cryptocurrency market return and sentiment. A bi-directional cascading effect infers a sentiment-driven bubble during the bubble period, whereas a reversal effect after the bubble implies a correction for overreaction. Indeed, the results are robust after controlling for market microstructure noise following Chordia et al. (2001). Out-of-sample tests further reveal that investor sentiment conveys incremental predictability relative to the benchmark strategy, especially during a non-bubble period. A trading strategy based on investor sentiment generates a much higher portfolio return (28 daily bps) than the CRIX return (19 daily bps).

The paper is structured as follows. Section 2 describes the data. Section 3 presents the methodology and compares the accuracy of different lexicon methods used for textual sentiment analysis. Section 4 describes the local-momentum autoregression model and Section 5 presents the empirical results. Section 6 provides robustness checks. Section 7 concludes.

2 Data

2.1 CRIX

The CRIX is chosen to represent the entire cryptocurrency market. The CRIX (CRyptocurrency IndeX) is created by Trimborn and Härdle (2018) and used to track the entire cryptocurrency market performance as close as possible. It is constructed robustly in the sense it considers a frequently changing market structure, hence the representativity and the tracking performance can be assured. In such a way, the number of constituents is changing over time, depending on market conditions and the relative dominance among cryptos. The data series starting from July 2014 can be downloaded through `thecrix.de`. The reallocation of the CRIX happens on a monthly and quarterly basis. It adopts a liquidity rule when incorporating a certain cryptocurrency into CRIX, and hence guarantees the trading of CRIX, which is good for ETFs and traders. CRIX has been widely investigated in the pioneering research on cryptocurrencies, including Hafner (2018), Chen et al. (2018), and da Gama Silva et al. (2019).

2.2 StockTwits

StockTwits[3] is a social microblogging platform similar to Twitter, but dedicated to financial discussion. Individuals, investors, market professionals, and companies can express opinions, spread news, advertise, etc., by posting messages with a maximum length of 140 characters. According to StockTwits, more than one million users now use the platform to share information and ideas, reaching an audience of more than 40 million people across the financial web and social media. Conversations are organized around "cashtags" (e.g. $SPY for S&P 500) that allows to narrow streams down to specific assets. Users can also express their sentiment by labeling their messages as "Bearish (negative) or Bullish (positive) *via* a toggle button.

New cryptocurrencies are regularly added to the list of cashtags supported by StockTwits.[4] A cashtag refers to a cryptocurrency if and only if it ends with ".X" (e.g. $BTC.X for Bitcoin, $LTC.X for Litecoin). We use this convention and StockTwits Application Programming Interface (API) to download all messages containing a cashtag referring to a cryptocurrency. StockTwits API also provides for each message its user's unique identifier, the time it was posted at with a one-second precision, and the sentiment associated by the user ("Bullish", "Bearish" or unclassified). Our final dataset contains 1,533,975 messages from 38,812 distinct users, posted between March 2013 and December 2018, and related to 465 cryptocurrencies (see Table 1 for a sublist of the collected currencies). Overall, 576,350 messages are classified as bullish (37.5%) and 130,511 as bearish (8.5%), and the remaining are unclassified. The imbalance between the numbers of positive and negative messages shows that online investors are optimistic on average, as previously found by Kim and Kim (2014) or Avery et al. (2016).

Figure 1 represents the number of messages per week related to cryptocurrencies on StockTwits, and CRIX weekly average. Investor attention has skyrocketed just like the prices did during the 2017 booming of the market. This indicates a certain relationship between investors behavior on StockTwits and price evolution.

Figure 2 represents the median, 5% and 95% quantiles of the volume of messages related to cryptocurrencies per hour, over several consecutive weeks, revealing strong intraday seasonality. Whereas global activity on StockTwits usually quiets down on weekends and outside market

[3]https://stocktwits.com/

[4]This list can be found at https://api.stocktwits.com/symbol-sync/symbols.csv.

Cryptocurrency	Message volume
Bitcoin BTC/USD ($BTC.X)	459,133
Litecoin ($LTC.X)	180,356
Tronix ($TRX.X)	162,429
Ripple ($XRP.X)	126,053
Ethereum ($ETH.X)	122,620
Verge ($XVG.X)	69,066
Bitcoin Cash ($BCH.X)	26,886
NEO ($NEO.X)	18,733
Stellar Lumens ($XLM.X)	18,196
Cardano ($ADA.X)	15,758
IOTA ($IOT.X)	15,347
NewYorkCoin ($NYC.X)	13,301
Po.et ($POE.X)	9,335
ReddCoin ($RDD.X)	9,316
PacCoin ($PAC.X)	6,121
Vechain ($VEN.X)	5,461
EOS ($EOS.X)	5,064
Binance ($BNB.X)	4,871
Monero ($XMR.X)	4,536
EthLend ($LEND.X)	4,246
ICON Project ($ICX.X)	3,811
Ethereum Classic ($ETC.X)	3,802
IOStoken ($IOST.X)	3,798
FunFair ($FUN.X)	3,627
Siacoin ($SC.X)	3,610

Table 1: 25 biggest cryptocurrencies available on StockTwits by message volume (up to 2017-12-31)

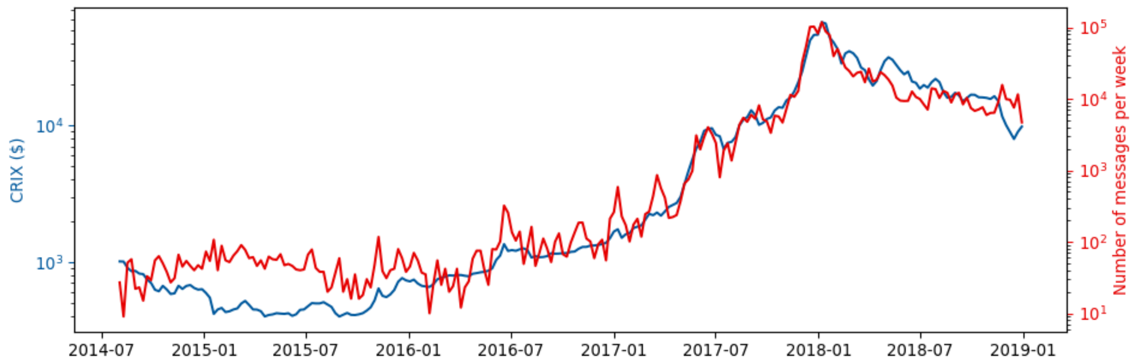

Figure 1: Message volumes and crypto market index
This figure presents the weekly number of crypto-related messages on StockTwits (in red) and the CRIX value (in blue, log scale) between June 2014 and December 2018

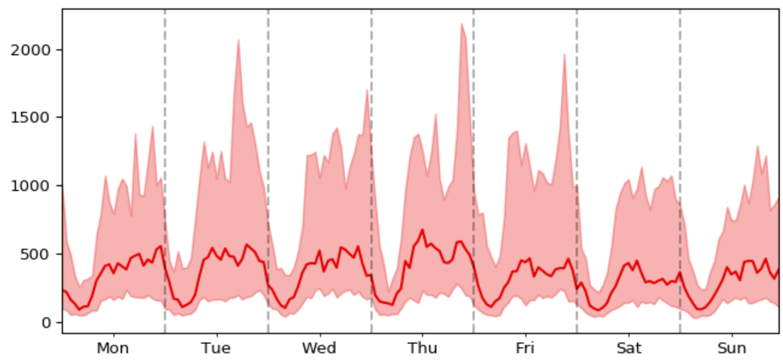

Figure 2: Number of messages per hour interval

This figure presents the median number of messages per hour related to cryptocurrencies posted on StockTwits between November 2017 and December 2017, and 5% and 95% quantiles. Dashed vertical lines indicate 00:00 UTC-05:00.

opening hours (Renault, 2017), crypto-related message volume is approximately the same every day, and only falls at night. Since the crypto market is open 24/7, it is natural to have activity on weekends too. The drops we observe at night indicate StockTwits is mostly used by people living in the US, which could constitute a sampling bias since the crypto market is worldwide.

2.3 Reddit

Reddit[5] is a discussion website where users can hold conversations by posting messages. On Reddit, posts are organized by topics called "subreddits". Using the Python Reddit API Wrapper (PRAW), we extract all messages posted on the eight subreddits with the highest number of subscribers: "CryptoCurrency","CryptoCurrencyTrading","CryptoMarkets", "Bitcoin", "BitcoinMarkets", "btc", "ethereum", "ethtrader".[6] For each message, we extract the exact date at which the message was posted, the author of the message, the content of the message and the number of upvotes/downvotes. We end up with a database of 1,392,587 messages posted between January 2014 and August 2018. Data from Reddit are in a sense pretty similar to data from Yahoo! Finance or Raging Bulls messages boards used by Antweiler and Frank (2004) and Das and Chen (2007). We choose to use data from Reddit as it is now by far the most visited message boards in the world. According to the last statistics from Alexa, Reddit is now the sixth most visited website in the world, just after Google, Youtube, Facebook, Baidu

[5] https://www.reddit.com/

[6] See https://www.reddit.com/r/CryptoCurrency/ for an example of the "CryptoCurrency" subreddit

and Wikipedia.[7] Furthermore, Reddit is a generic message board, and not a message board only dedicated to financial markets, allowing us to capture a wider number of topics related to cryptocurrencies including discussions about cryptocurrency technologies and the blockchain.

3 Textual sentiment analysis

To test the forecasting power of user-generated sentiment related to cryptocurrencies, we need to develop a methodology that converts text data into a quantitative sentiment variable. Two main strategies can be used for this purpose: dictionary-based and machine learning-based analysis. In this section, we justify why we only use the first one, detail our methodology and compare our results to benchmark dictionaries used in the literature.

3.1 Dictionary-based analysis

A dictionary, or lexicon, is a list of words labeled as positive, negative or neutral. Assuming such a list, the classic "bag-of-words" approach consists of counting the number of positive and negative words in a document to assign a sentiment value, or tone, to it. For example, a simple dictionary containing only the words "good" and "bad" with respectively positive and negative labels would classify the sentence "Bitcoin is a good investment" as positive, with a tone of +1.

The construction of the lexicon itself can be achieved in several ways. One technique is to have experts (researchers or practitioners) list and classify words they estimate to be meaningful, based on their knowledge. Another approach is to automatically select recurrent words from a set of documents, and then to have experts classify them. This method has been used by Loughran and McDonald (2011) to create their financial dictionary (from now, referred to as LM). Finally, one can also implement a fully automatic procedure extracting relevant words from a set of documents and computing statistical measures based on their frequency to classify them as positive, negative or neutral. Renault (2017) proceeds this way to construct his social media lexicon (from now, referred to as GL).

The simplicity of the dictionary-based approach guarantees transparency and replicability

[7] https://www.alexa.com/topsites

provided the lexicon is made public. However, this also comes with limitations associated with two main challenges of natural language analysis. First, language interpretation is sensitive to the "context of a discourse" (Deng et al., 2017). For example, Loughran and McDonald (2011) point that words like "tax" or "cost" are classified as negative by the Harvard General Inquirer Diction, whereas they should be considered as neutral in a financial context. The second challenge is related to the "lexical and syntactical choices of language" Deng et al. (2017). One example would be the difference between newspapers where a formal and standardized tone is mostly used, and social media, where slang and emojis are preponderant (Loughran and McDonald, 2016). It is thus necessary to develop and use appropriate context-specific lexicons. Another arguable feature of classic lexicons is they only have two categories (positive and negative), somehow assuming words all have the same explanatory power. Using term weighting instead of term classification can remedy it, but also creates a risk of over-fitting as many weighting schemes can be chosen. Finally, a bag-of-words model is insensitive to word order and grammar and thus has a weak understanding of language structure. One solution to this is to consider n-grams (i.e., terms in the dictionary that are sequences of n words). For example, including the bigram "not good" as negative probably avoids some misclassifications.

While it is possible to refine the dictionary model to address some of its issues, another approach is to turn to machine learning techniques. These are statistical algorithms that need to be trained on pre-classified documents before they can classify themselves new documents. In Renault (2017), the author implements an automatic procedure for building a term-weighting social media lexicon that includes bigrams and compares its accuracy to a machine learning classifier and several benchmark lexicons on StockTwits messages data. This results in high, but similar, accuracy improvement of the first two over the benchmarks. As we work on similar data, better transparency and close accuracy justify our choice to only focus on dictionary analysis.

3.2 Building a crypto-specific lexicon

We follow Oliveira et al. (2016) automated procedure to build our lexicon from StockTwits messages, which is possible thanks to two features of StockTwits. First, messages contain explicit reference to the asset they mention *via* the "cashtag" system, which allows us to select

Before processing	After processing
$BTC.X why can’t it hold $14k ?? Shameless pumpers said 25 by Christmas 😩	cashtag why can t it hold moneytag ? ? shameless pumpers said numbertag by christmas 😩
$BTC.X Merry Xmas to all coiners and no coiners alike! 2018 is gonna be lit!! 💰🎄🚀	cashtag merry xmas to all coiners and negtag_coiners alike ! numbertag is gonna be lit ! ! 💰 🎄 🚀
$XVG.X all greeeeeeeeb 😂😂	cashtag all greeeb 😂 😂
$NEO.X In NEO I trust!!! https://neousd.bid/	cashtag in neo i trust ! ! ! linktag

Table 2: Pre-processing of StockTwits messages

messages that only refer to cryptocurrencies. Second, users are able to label their message as "Bullish" (positive) or "Bearish" (negative) when they post it, which provides us with a large training dataset adapted to supervised learning. We consider all messages labeled as "Bullish", and randomly split it into a training positive dataset (75% of all positive messages, i.e. 432,262) and a testing positive dataset (the remaining 25%, i.e. 144,087). To avoid domination of the corpus by excessively prolific users (possibly robots), we impose a maximum proportion of 1% of the dataset per user, as in Pang et al. (2002). Proceeding identically with negative messages, we constitute our final training and testing dataset.

Our natural language processing methodology is inspired by Sprenger et al. (2014) and Renault (2017). First, all messages are lowercased. To account for lengthening of words, which is a critical feature of sentiment expression on microblogs (Brody and Diakopoulos, 2011), sequences of repeated letters are shrunk to a maximum length of 3. Tickers ("$BTC.X", "$LTC.X"...), dollar or euro values, hyperlinks, numbers and mentions of users are respectively replaced by the words "cashtag", "moneytag", "linktag", "numbertag" and "usertag". The prefix "negtag_" is added to any word consecutive to "not", "no", "none", "neither", "never" or "nobody". Finally, the three stopwords "the", "a", "an" and all punctuation except the characters "?" and "!" are removed. Exclamation and interrogation marks are kept as it has been previously shown that they are often part of significant bigrams that improve lexicon accuracy (Renault, 2017), which is confirmed by our findings (see Table 3 below). Table 2 shows examples of messages before and after processing.

We then compute for each term t found in our dataset (t can either be a unigram or a

Term	*SW*	Positive	Negative
🚀	0.97	15313	31
bears	0.59	5878	291
hodl	0.32	4551	441
hodl !	0.64	632	26
dump	-0.75	681	937
💩	-0.91	114	519
moon	0.45	2199	160
🔥	0.40	1991	164
💰	0.75	2359	64
binance	0.61	1892	88
💩 💩	-0.98	15	303
market cap	0.36	937	83
cashtag hodl	0.32	870	86
😎	0.52	1009	61
on coinbase	0.31	772	77
shitcoin	-0.88	50	151
numbertag coins	0.62	827	37
💪	0.79	909	20
wallet	0.56	735	39
bittrex	0.76	852	22
alt	0.31	571	57
📈	0.75	803	22
btfd	0.42	599	47
tulip	-0.83	57	122
ico	0.40	580	47
📉	-0.91	30	123
amazon	0.47	362	25
tron	0.61	397	18
tulip mania	-0.94	9	58
🌙	0.94	395	2
millennials	-0.76	36	51
scam coin	-0.86	19	51
drug	-0.76	26	38
buttcoin	-0.95	5	39
hackers	-0.83	15	33
cashtag fomo	0.33	135	13

Table 3: Frequent crypto-specific terms in the lexicon. *SW* is the sentiment weight computed as in (1). A complete list of crypto-specific terms, the frequency of appearing in the bullish and bearish messages, and the corresponding sentiment weights can be found in https://sites.google.com/site/professorcathychen/home.

bigram) a sentiment weight $SW(t)$ defined by:

$$SW(t) = \frac{f_{\text{pos}}(t) - f_{\text{neg}}(t)}{f_{\text{pos}}(t) + f_{\text{neg}}(t)}, \tag{1}$$

where $f_{\text{pos}}(t)$ ($f_{\text{neg}}(t)$) is the term-frequency of t in the positive (negative) corpus, i.e.:

$$f_{\text{pos}}(t) = \frac{\text{Nb. of occurrences of } t \text{ in positive messages}}{\text{Nb. of terms in positive messages}} \tag{2}$$

Finally, we sort all terms by their sentiment weight and only retain the first and last quintiles.

Figure 3: Word clouds of crypto-specific terms with positive (left cloud) or negative (right cloud) sentiment weights.

Table 3 shows a list of selected crypto-specific terms, namely terms that appear exclusively in our lexicon but not in LM or GL.[8] Figure 3 shows word clouds of crypto-specific terms with a positive (negative) sentiment weight. One can find several crypto-related terms such as "hodl" (a misspelling of "hold" that has become popular among crypto-investors), "shitcoin" and "buttcoin" (puns usually used to designate small, alternative coins), or "ico" (Initial Coin Offering, a means of crowdfunding the release of a new cryptocurrency). The lexicon also contains a large amount of emojis that are not crypto-related by themselves but appear as key features of the crypto-investors language. Finally, we can observe bigrams do provide additional information: for example "hodl !" has a sentiment weight of 0.64 while "hodl" has a sentiment weight of 0.32. Note that in Table 3 the last two columns list the count of the crypto-specific term appearing in the positive and negative messages, respectively. These two columns need to be transformed into the frequency measure in (2) to calculate the SW value.

3.3 Classification accuracy

We use the testing dataset constructed in section 3.2 to compute the out-of-sample accuracy of the three following lexicons:

- CL (Cryptocurrency-specific Lexicon) - weighted lexicon with approximately 4,800 positive and 4,800 negative terms
- GL (General Lexicon for social media) - weighted lexicon with approximately 4,000 positive and 4,000 negative terms,

[8]A complete list of crypto-specific terms, the frequency of appearing in the bullish and bearish messages, and the corresponding sentiment weights can be found in https://sites.google.com/site/professorcathychen/home

Lexicon	CC	CC_{bull}	CC_{bear}	CM	CM_{bull}	CM_{bear}
CL	0.86	0.89	0.70	0.83	0.83	0.83
GL	0.72	0.71	0.80	0.79	0.78	0.82
LM	0.54	0.51	0.68	0.22	0.21	0.23

Table 4: Classification accuracy of several lexicons. Includes, for each lexicon, the proportion of correct classifications (CC) among all classified messages, and per class (CC_{bull} and CC_{bear}); and the proportion of classified messages (i.e. containing at least one of the terms of the lexicon), overall and per class (CM, CM_{bull} and CM_{bear}).

- LM (Loughran-McDonald Lexicon) - 354 positive terms and 2,355 negative terms, with SW equal to 1 (resp. -1) for positive (resp. negative) terms.

For each message in the testing dataset, its tone is defined as the mean sentiment weight of its terms. It is then compared to the sentiment declared by the user who posted the message. If the tone is 0 (or no terms in the message belong to the lexicon being evaluated), the message is considered unclassified; if the tone is positive (resp. negative) and the message is declared bullish (resp. bearish), it is considered correctly classified. We then compute for each lexicon the proportion of correct classifications (CC) among all classified messages (CM), the proportion of classified messages among all messages, the proportions of correct classifications per class (CC_{bull} and CC_{bear}), and the proportions of classified messages per class (CM_{bull} and CM_{bear}). Results are reported in table 4.

We find that LM performs badly at classifying crypto messages, with only 22% of messages being classified and 54% of correct classifications among them. The accuracy of the GL of social media is better, with 79% of classified messages, and 72% of correct classifications, confirming that the language domain on social media has to be handled with specific tools. Finally, CL manages to classify 83% of messages, with 86% of them being correctly classified, which is a 20% improvement over GL. This result confirms the necessity of using a specific lexicon to measure sentiment on a specific market.[9]

[9] To facilitate replicability and facilitate further research in the area, our crypto-specific lexicon will be available online.

3.4 Investors' Sentiment Index

One might define investor sentiment as optimism or pessimism about stocks in general (Baker and Wurgler, 2006). Sentiment quantification and construction in the classical financial asset classes is almost well-established and well-studied. However, this issue in the new digital asset class is not fully discovered yet. We, therefore, aim to construct an aggregated cryptocurrency sentiment index that encompasses massive opinions from granular users in social media forums.

Throughout the paper, we focus on the StockTwits messages as the source of sentiment measures because the messages there are prone to reflect investors' sentiment concerning financial perspectives such as trading behavior, risk preference, and speculation. Reddit messages are often linked to technological discussions and employed for the robustness tests. Later on in Section 5 we analyze the information content of the sentiment measures from these two message sources.

The individual message sentiment score is defined as the average $SW(t)$ of the terms present in the message. We use our lexicon (CL) and other two benchmark lexicon (GL and LM) to derive a sentiment score between 1 and +1 for all 1,533,975 messages posted on StockTwits between March 2013 and December 2018, and for the messages in Reddit, respectively. Then, we compute daily investor sentiment indicators by averaging, at 24-hour intervals, the sentiment score of individual messages published per calendar day. We denote those measures $\mathrm{S}_t^{type,source}$ where *type* $\in \{CL, GL, LM\}$ and source $\in \{StockTwits, Reddit\}$.

It is worth noting that CL-based sentiment and GL-based sentiment seem to be closely related but are still different from each other. Meanwhile, both of them deviate significantly from LM-based sentiment. The correlation between $S^{CL,StockTwits}$ and $S^{GL,StockTwits}$ is 0.83 while it is merely 0.37 between $S^{CL,StockTwits}$ and $S^{LM,StockTwits}$. This further supports the need of constructing a dictionary for a crypto asset class.

Table 5 reports summary statistics for the main variables used in this paper. The daily excess CRIX return has a mean of 0.21% and a standard deviation of 3.86%, implying a monthly Sharpe ratio of 0.30. Interestingly, the CL-based index is generally positive whereas other types of indices using GL and LM show negative opinions, indicating that CL-based measures are relatively more optimistic than GL and LM. This can be attributed to the fact that CL captures more domain-specific terms exclusively for cryptocurrencies that are not well covered by GL and LM list, such as "hodl" or 🚀. As a consequence, the optimism might be underestimated in the classical lexicon.[10]

Variable	Mean	Std.	Skew.	Kurt.	Min.	Max.	$\rho(1)$
$R_{m,t}$	0.0021	0.0386	-0.2869	8.3774	-0.2238	0.2196	-0.0169
$S^{CL,StockTwitts}$	0.2201	0.165	-0.9652	3.1946	-0.2548	0.5874	0.9647
$S^{GL,StockTwitts}$	-0.0397	0.116	0.2728	3.7572	-0.3495	0.3893	0.9283
$S^{LM,StockTwitts}$	-0.036	0.3341	0.4083	4.2145	-1	1	0.9123
$S^{CL,Reddit}$	0.2208	0.0416	-0.9776	5.9879	0.0000	0.3188	0.9748
$S^{GL,Reddit}$	-0.0175	0.0262	0.3568	2.9725	-0.0853	0.0587	0.9871
$S^{LM,Reddit}$	-0.0361	0.0225	-1.0054	5.5332	-0.1450	0.0247	0.9622
$\text{MsgVol}_{StockTwits}$	4.0513	2.6425	0.4674	1.8048	0	10.1589	0.9663
MsgVol_{Reddit}	6.3130	0.8104	0.9284	2.9146	4.7622	8.8778	0.9590

Table 5: Summary Statistics

This table reports summary statistics for the CRIX return ($R_{m,t}$), investors' sentiment measures and the logarithm of message volume + 1 denoted as MsgVol. CRIX data is collected from http://data.thecrix.de/ constructed by Trimborn and Härdle (2018), message data is collected from StockTwits and Reddit. The sample period of Reddit message is from 2014-08-01 to 2018-06-30 and the other variables is from 2014-08-01 to 2018-12-27.

4 Bubble and fundamental value

Market efficient hypothesis and the asset pricing theory lie on the postulation that asset prices move around the trajectory of fundamental value. In other words, asset prices in the long-run should mean revert to their corresponding latent fundamental value. However, the presence of a bubble indicates a locally explosive autoregression process and a salient deviation from the fundamental value. While investor sentiment should only have a short-lived effect if prices tend to converge to their fundamental value, its impact might be prolonged during a bubble period whenever the market fails to correct itself.

The cryptocurrency market had experienced a striking bubble from 2017 until Jan. 2018, which is conjecturally caused by sentiment (Bukovina et al., 2016). We examine the tendency of converging to the fundamental value in the long-run *before* and *post*-bubble period, respectively. We, therefore, employ the local-momentum autoregressive model proposed by Duan (2016) to characterize a parsimonious autoregressive model in the time series of CRIX that is globally mean-reverting into its fundamental process but locally driven by momentum. This local-

[10]Another clue can be referred to Table 4 where crypto-specific lexicon shows a remarkable classification accuracy in comparison with the other two.

momentum autoregressive process comprises two parts, central tendency concerning the fundamental value process and stochastic deviation on its own. The model permits a mean-reverting behavior on a larger timescale while it displays a local momentum on a smaller timescale. A local-momentum autoregressive process therefore concurrently exhibits stochastic central tendency and local momentum, which applies to the specification of asset dynamics in practice.

The co-existence of long-run mean-reverting and short-run momentum is very likely to be observed in crypto assets. During the bubble period, a local momentum becomes throughout pervasive which ultimately deters a tendency from the central value. In the research of De Long et al. (1990b), a momentum trading or positive feedback trading induced by the sentiment of noise traders is anticipated by rational speculators, hence rational speculators who trigger the noise traders' positive feedback trading destabilize the market in the short-run but stabilize it in the longer time frame. The econometric model by Duan (2016) is very ideal for econometrically testing the financial model by De Long et al. (1990b) in which it generates a positive correlation of stock returns at short horizons (characterized by a momentum) as positive feedback traders respond to past price increases, and negative correlations of stock returns at long horizons, as prices eventually revert to fundamentals (characterized by a central tendency factor).

4.1 Local-momentum autoregressive model

Suppose we have a series of log prices for the CRIX, denoted X_t. X_t in the framework of local-momentum central tendency autoregression (LM-CTAR) model is specified as follows:

$$\Delta X_t = \kappa_x(\mu_t - X_{t-1}) + \omega(\bar{X}_{(t-1)|n} - X_{t-1}) + \sigma_x \varepsilon_t \tag{3}$$

$$\Delta \mu_t = \kappa_\mu(\bar{\mu} - \mu_{t-1}) + \sigma_\mu \epsilon_t \tag{4}$$

$$\bar{X}_{(t-1)|n} = \sum_{i=t-n}^{t-1} b_{t-i} X_i \tag{5}$$

where $\Delta X_t = X_t - X_{t-1}$, ε_t and ϵ_t follow standard normal distribution and are independent, conditional on the filtration generated by $\{X_s : s \leq t\}$. $\sigma_x > 0$, $\sigma_\mu > 0$, $\kappa_x \geq 0$, and $\sum_{i=1}^{n} b_i = 1$ with $b_i \geq 0$. Here we take $b_i = 1/n$, a simple equally weighted scheme. It is well known that under $0 < \kappa_x < 2$, X_t is a strictly stationary and ergodic process. $\kappa_x = 0$ leaves us a unit-root process. In this case, X_t is very hardly converging to its central tendency factor, μ_t, a stochastic process being characterized as a mean-reverting process with $\bar{\mu}$ denoted as long-run level of fundamental value and κ_μ governing the speed of reversion. The LM-CTAR model in (3) postulates that an observed price stochastic process moves toward a latent fundamental stochastic process in the long run. A mean-reverting alike stochastic process for fundamental factors can be found in Balduzzi et al. (1998) and Peng (2005).

One can express the above model in a matrix-vector form:

$$\mathbf{X}_t = \mathbf{A} + \mathbf{B}\mathbf{X}_{t-1} + \mathbf{Z}_t \tag{6}$$

where $\mathbf{X}_t$, $\mathbf{Z}_t$, $\mathbf{A}$ are n-dimensional column vectors, and $\mathbf{B}$ is an $n \times n$ matrix.

$$\mathbf{X_t} = \begin{bmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-n+1} \end{bmatrix} \mathbf{Z_t} = \begin{bmatrix} \kappa_x(\mu_t - \bar{\mu} + \sigma_x \epsilon_t) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \mathbf{A} = \begin{bmatrix} \kappa_x \bar{\mu} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \mathbf{B} = \begin{bmatrix} 1 - \kappa_x - \omega(1 - b_1) & \omega b_2 & \cdots & \omega b_n \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

The $\mathbf{B}$ coefficient matrix determines the stationarity of model, particularly from the first row of it. In the special case, if κ_x and ω are both equal to zero, the spectral radius of $\mathbf{B}$, denoted as $\rho(\mathbf{B})$, will be equal to one. The model is explosive with $\rho(\mathbf{B}) > 1$ while $\kappa_x = 0$ and $\omega < 0$. $\rho(\mathbf{B}) < 1$ ensures the stationarity.

The specification of local momentum lies on the second component of (3). It characterizes a momentum build up by the process' own past n-period moving average in a relative short-run time frame. $\omega > 0$ defines a *local momentum-preserving* feature, while $\omega < 0$ exhibits a *local momentum-building* characteristic. In the special case with $\kappa_x = 0$ and $\omega < 0$, X_t will behave a bubble-like price dynamics, akin to a locally explosive autoregressive process. The employed LM-CTAR model is generic enough to specialize in a variety of stochastic processes. A more detailed discussion can be found in Duan (2016). Note that the discrete-time version can be rewritten as a continuous-time framework if a continuous-time model is preferred in the study.

4.2 Bubble growing period versus post-bubble period

To investigate the tendency of CRIX time series reverting to its fundamental stochastic process and its local momentum behavior, in the bubble-growing and in the post-bubble period, we undertake the LM-CTAR model for the bubble-building period (Aug. 2014 - Jan. 2018) and for the post-bubble period (Feb. 2018 - Dec. 2018), respectively. This design ought to enhance our understanding of price dynamics under different phases where sentiment may have a different impact.

A growing body of literature has paid their attention to test cryptocurrency bubbles, either statistically or econometrically. Cryptocurrency markets, with the market characteristics such as short-sale constraints, decentralization, and blockchain transaction, are attractive to researchers. The bubble tests for cryptocurrencies proposed by Hafner (2018) and Geuder et al. (2018) indicate that there are no further price bubble indications after January 2018. We, therefore, partition our sample period along with the accordance between two studies.

Worth noting that the LM-CTAR process is a state-space model due to an inclusion of a latent process, μ_t. The estimation for such a process requires help from the Kalman Filter with a further assumption imposed for the innovations. The measurement is (3) whereas the transition equation is (4). The estimation results are reported in Table 6. Remarkably, the CRIX process behaves differently across two subperiods. During the bubble-growing period, the long-run mean reverting to fundamental factor is hampered given an insignificant κ_x, whereas it turns into moving toward the central tendency process after the bubble as κ_x becomes significant. Another striking evidence is that before, and also during the bubble forming, the latent central tendency process is almost impossible given the insignificant κ_μ, whereas the fundamental process emerges after the bubble. It turns out that during the post-bubble period cryptocurrency market is able to form a stationary fundamental process and push the price process close to the fundamental process. While comparing the magnitude of κ_x and κ_μ, the speed of stochastic fundamental process reverting to its long-run mean level is much slower than that of the price process. It is understandable since the fundamental process represents the long-run price process, it behaves less volatile ($\sigma_\mu = 0.0015 < \sigma_x = 0.0047$), and slowly converges to its long-run mean value.

The local momentum process also evolves differently across two subperiods, a momentum-preserving process while the bubble grows and a momentum-building pattern while the bubble bursts. An inference of a momentum-preserving ($\omega > 0$) or momentum-building ($\omega < 0$) hinges on the sign of estimated ω. The ω estimates are both significant but positive during the bubble growing and turning to be negative after the bubble. A local momentum-building-like price process implies a strong trend-chasing behavior. Investors sell more when prices drop more and a downward plunge triggers more sell, reflecting investors' fear of future price declines. Likewise, the pre-bubble and bubble period the market is characterized as a momentum-preserving behavior, in this situation, the trend over the past n days is preserved. Investors tracing the trend expect the trend to be kept in the short-run. A sharp price downward revision during the post-bubble period, documented by the momentum-building effect, indicates a likelihood of overreaction and a price overshoot trigger by investor sentiment. The rational speculators, in this case, switch their role from destabilizing the market by triggering the positive trading during the bubble to contributing the market stabilization in the post-bubble period by buying the undervalued coins. This change is anticipated since they encounter the limits to arbitrage in the former period.

The extant literature supports our findings for the local momentum discovery. Detzel et al. (2018) find that the trading strategy of Bitcoin traders emphasizes on the path of prices, a typically popular strategy for trend traders who utilize technical analysis. We use seven days to compute the moving average judging from the value of log-likelihood if the local momentum feature is presented. The estimate of ω remains barely changed and is robust when the window of moving average is fine-tuned to six or eight days. Regarding to stationarity, in Table 6, $\rho(\mathbf{B}) = 0.961$ during the bubble period is slightly higher than $\rho(\mathbf{B}) = 0.943$ after the bubble. Stationarity is slightly recovered after the bubble.

Having the estimates in Table 6, the distribution of the vector $(X_t, \mu_t)^\top$ conditional on X_{t-1} and μ_{t-1} and the covariance matrix of the measurement error, we can apply the standard Kalman filtering results to (4) to obtain the predicted mean and variance of the state variables.[11] Figure 4 depicts the predicted and the observed CRIX process. The predicted price series ties to the observed one, and both two in the longer time scale are prone to revert to the calibrated latent fundamental process which exhibits a very slow downward trend. A longer period is needed to see a mean-reverting feature in the fundamental process given its slow speed of

[11]More technical detail can be found in Appendix C of Duan (2016).

	Bubble-growing period		post-Bubble period	
κ_x	0.0460	(0.0376)	0.0438	(0.0152)
$\bar{\mu}$	0.0301	(0.0314)	0.8392	(0.1334)
σ_x	0.0036	(0.0001)	0.0047	(0.0003)
κ_μ	0.0001	(0.0005)	0.0028	(0.0011)
σ_μ	0.0064	(0.0015)	0.0015	(0.0006)
ω	0.0597	(0.0194)	-0.0611	(0.0034)
$\log L$	5346.42		1318.14	
$\rho(\mathbf{B})$	0.961		0.943	

Table 6: Estimation results of the LM-CTAR model during two sub-periods

Estimation results for the LM-CTAR model for the bubble-growing period (Aug. 2014 - Jan. 2018, 1280 observations) and for the post-bubble period (Feb. 2018 - Dec. 2018, 344 observations), respectively. Standard errors are in parentheses. $\log L$ is the value of the log likelihood function. $\rho(\mathbf{B})$ is the spectral radius of autoregressive coefficient matrix, $\mathbf{B}$, used to determine the stationarity of the LM-CTAR process. The estimation is done by the Kalman Filter algorithm and the price series has been taken logarithm first and multiplied 0.1.

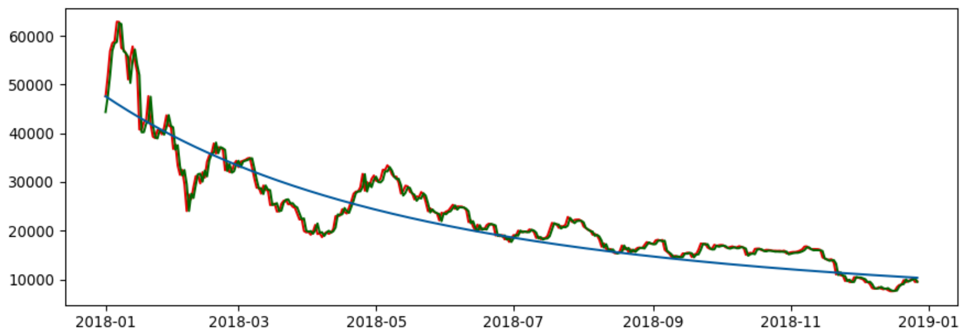

Figure 4: The evolution of price (red), predicted price (green) and inferred fundamental value (blue) process after bubble

reversion. However, the reversion in the price process relative to the fundamental process is much quicker. For example, in 7th Feb. 2018, the price is below to its fundamental level, and a reversion is anticipated and observed.[12] Likewise, the correction for the overvalued price is evident in May, Aug., Oct. and Nov. 2018. It's predicted that, due to the limits-to-arbitrage, the duration to revert to the fundamental value in the overvalued case is longer than in the undervalued case, reflecting a slower speed of reversion in the overvalued case. With such a scheme, one can form the arbitrage strategies to take the arbitrage opportunity through a temporal deviation between the price and fundamental process.

[12]Later, in Section 5, we examine that a temporal deviation can be driven by sentiment-trading or momentum trading.

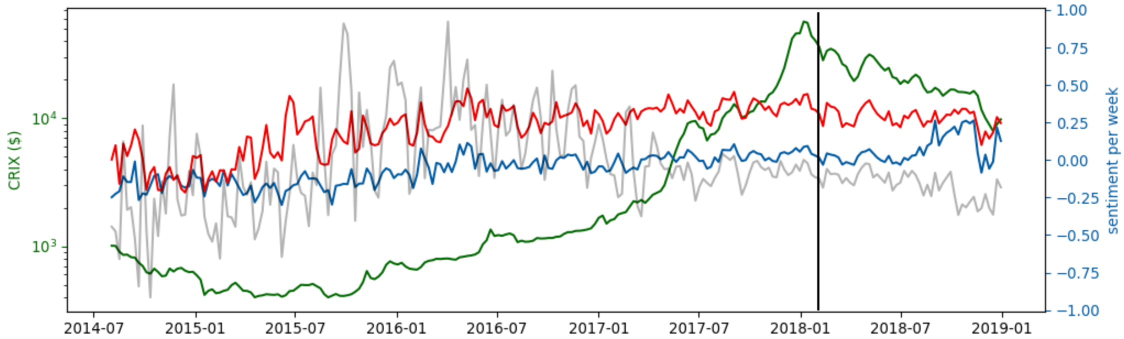

Figure 5: The evolution of CRIX (green), sentiment series constructed by CL (red), GL (blue) and LM (grey) in a weekly frequency. The straight line separates the subperiods based on the bubble test by Hafner (2018).

5 Bubble, sentiment and cryptocurrency return

A bubble is statistically characterized as an explosive autoregressive process and is evident economically by a prolonged deviation from its fundamental value. In this section, we examine how bubble formation and sentiment-based trading are related to each other. To explore this relation, the cryptocurrency market has a clear advantage over stock markets as intrinsically it is unregulated, decentralized and has a hard-to-measure fundamental value. Since long-term fair value and prospects are relatively intangible, and, more importantly, since this market lacks the institutional participants due to the regulatory constraints, trading position taken by sentiment-driven traders may lead to the formation of a bubble. We quantify their sentiment by analyzing their conversations and messages in social media and by using our state-of-art lexicon to precisely capture their moods and opinions.

To answer the research question of whether there is a sentiment-driven bubble, we employ the vector autoregression (VAR) model for a bidirectional feedback specification. Figure 5 displays the interplay between CRIX series (in the logarithm level) and three types of sentiment index using the StockTwits messages described in section 3.4. One can observe that the CRIX series comove mainly with the CL-based one and moderately with the GL-based one, but not with LM-based sentiment. The cryptocurrency market had experienced large shocks, including the shutdown of Mt. Gox in 2014 and a big soar in mid of 2017 until Jan. 2018, and a crash since Feb. 2018. During the investigating period, the interaction between CRIX return and sentiment seems subject to the life cycle of a bubble.

The vector autoregression (VAR) model is simple and has been widely applied across many topics in finance, especially for an endogenous system problem. It is conjectured that cryptocurrency return and the sentiment observed in the social media influence each other mutually and intertemporally. More specifically, an optimistic sentiment causes a price soar, and a subsequent positive-feedback or momentum trading given price soars inflates investor sentiment further.

A $p-$th order VAR, denoted as VAR(p), is described as:

$$\mathbf{Y}_t = c + \mathbf{A}_1\mathbf{Y}_{t-1} + \mathbf{A}_2\mathbf{Y}_{t-2} + \cdots + \mathbf{A}_p\mathbf{Y}_{t-p} + e_t \tag{7}$$

where $\mathbf{Y}_t = [R_t, S_t]^\top$, R_t is the log return of CRIX at time t, S_t is the sentiment measure at time t described in section 3.4. The CL-based sentiment is employed due to its validated accuracy in Table 4. c is a 2-vector of constants (intercepts), $\mathbf{A}_i$ is a time-invariant (2×2) -matrix and e_t is a 2-vector of error terms satisfying the conditions such as zero mean, no serial correlation. $p = 3$ is suggested by the goodness of fit criteria, Bayesian information criterion.

Table 7 reports the estimates of the first autoregression process on R_t, while Table 8 shows the second autoregression process on S_t. Both processes are estimated under a sequence of its lagged terms and the lagged other variable up to $t-3$. For a sake of brevity and for emphasis on a mutual feedback effect, in both tables, we suppress the display of the estimates from its own lagged term.[13] As such, the autoregression coefficients tagged to $S_{t-1}, S_{t-2}, S_{t-3}$ in Table 7 are selectively reported due to a focal discovery on the reversal effect caused by behavioral overreaction. For the purpose of comparison, we have considered all the sentiment measures defined in the previous section in the VAR exercise. However, as the reversal and cascading effect become weaker when using the GL-based sentiment and vanish in the LM-based case, we choose to present our results only for the CL-based sentiment indicator.

During the bubble period, the positive sign and the significance of S_{t-1} on R_t indicates a sentiment-driven price increase. In Table 8, we also observe an increasing price pumping up sentimental exuberance. A positive feedback or momentum trading strategy becomes prevalent as traders buy, given their exuberant sentiment, after observing a price soar. The resulting cascading effect through a bidirectional feedback system makes the bubble growing. This drawn conclusion is, however, exclusive for the sentiment distilled from StockTwits messages. The

[13]The complete results can be provided by request.

	Bubble-growing period		post-Bubble period	
R_t	StockTwits	Reddit	StockTwits	Reddit
S_{t-1}	0.0474	0.0145	0.4793	0.4317
	(0.050)	(0.679)	(0.011)	(0.008)
S_{t-2}	0.0028	-0.0109	-0.1495	-0.4754
	(0.935)	(0.765)	(0.023)	(0.007)
S_{t-3}	-0.0318	0.0266	-0.2639	0.0779
	(0.192)	(0.439)	(0.000)	(0.634)
adjustR^2	0.014	0.009	0.241	0.053

Table 7: Feedback effects from return to sentiment conditional on bubble periods
This table documents a mutual feedback system constituted by the CRIX return (R_t) and sentiment measure (S_t), for the bubble-growing period (Aug. 2014 - Jan. 2018, 1280 observations) and for the post-bubble period (Feb. 2018 - Dec. 2018, 344 observations for stocktwits; Feb-Jul. 2018, 181 observations for reddits), respectively. p-values are in parentheses. The feedback of R_t on the lagged S_t is particularly reported. For a sake of brevity, the estimates from its own lagged term are suppressed.

extracted opinions from Reddit simply reflects the existing price soar, but hardly drive the future price movement.

While turning to the post-bubble period, a cascading effect has broken down, resulting in a bubble burst. Investor sentiment is no longer being governed by the previous price change shown in Table 8, impeding from a positive feedback trading. Nevertheless, the sentiment at $t-1$ continues to drive price movement, however, the observed price reversal at $t-2$ and $t-3$ is the consequence of price correction for a behavioral overreaction.

The empirical evidence in Table 7 and 8 elaborately confirms the model of positive feedback trading proposed by De Long et al. (1990b). In De Long et al. (1990b), the presence of bubble relies much on the positive feedback trading strategy implemented by a significant amount of investors, aggravated by arbitrageurs' anticipatory pumping up the bubble. For arbitrageurs, social media is a good venue to form their anticipation and to rule the market sentiment further. As pointed by De Long et al. (1990b), *purchases by rational speculators can make positive feedback traders even more excited and so move prices even further away from fundamental values than they would reach in the absence of rational speculators.* As this excitement gradually declines to reflect a price stumble, rational arbitragers anticipate that sentiment momentum is no longer sustainable, indicating the timing of collapses of the bubble. They eventually act as speculators by pushing the price in the direction of fundamental value.

	Bubble-growing period		post-Bubble period	
S_t	StockTwits	Reddit	StockTwits	Reddit
R_{t-1}	0.0980	0.1101	0.0606	-0.0033
	(0.002)	(0.002)	(0.295)	(0.930)
R_{t-2}	0.0107	0.0292	0.1555	0.0829
	(0.738)	(0.204)	(0.004)	(0.027)
R_{t-3}	0.0356	-0.0104	0.1186	-0.0477
	(0.270)	(0.648)	(0.030)	(0.207)
adjustR^2	0.943	0.448	0.685	0.250

Table 8: Feedback effects from return to sentiment conditional on bubble periods
This table documents a mutual feedback system constituted by the CRIX return (R_t) and sentiment measure (S_t), for the bubble-growing period (Aug. 2014 - Jan. 2018, 1280 observations) and for the post-bubble period (Feb. 2018 - Dec. 2018, 344 observations for stocktwits; Feb-Jul. 2018, 181 observations for reddits), respectively. p-values are in parentheses. The feedback of S_t on the lagged R_t is particularly reported. For a sake of brevity, the estimates from its own lagged term are suppressed.

6 Robustness check

This section provides supplementary analyses for short-run return predictability. Alternative predictors are first controlled in the VAR regression model as exogenous variables to remove market microstructure noises, and the out-of-sample forecast ability is subsequently evaluated and compared. The economic significance can be verified by implementing a trading strategy.

6.1 Market Microstructure Effect

Chordia et al. (2001) find that daily changes in market averages of liquidity and trading activity are time-varying and negatively autocorrelated. When stock returns decline, so does liquidity. Periods of volatility are followed by a decrease in trading activity. They also document day-of-the-week patterns, with Fridays experiencing lower trading activity and liquidity. In this case, our empirical results should take into account those market microstructure noises. To address these issues, we further control lagged volatility, lagged message volume and day-of-the-week effects in the VAR regression model. Table 9 and Table 10 report estimation results. Similar to Table 7 and Table 8, the sentiment effects under a bubble and post-bubble period remain after accounting for all potential microstructure factors.

	Bubble-growing period		post-Bubble period	
R_t	StockTwits	Reddit	StockTwits	Reddit
S_{t-1}	0.0498	0.0930	0.1557	0.0680
	(0.044)	(0.017)	(0.019)	(0.060)
S_{t-2}	-0.0415	-0.0915	-0.2031	-0.0700
	(0.237)	(0.101)	(0.008)	(0.088)
S_{t-3}	0.0020	0.0035	-0.0158	-0.0080
	(0.912)	(0.926)	(0.808)	(0.824)
adjustR^2	0.006	0.007	0.030	0.016

Table 9: Feedback effects from sentiment to return with market microstructure effects

This table documents a mutual feedback system (VAR regression) constituted by the CRIX return (R_t) and sentiment measure (S_t), for the bubble-growing period (Aug. 2014 - Jan. 2018, 1280 observations) and for the post-bubble period (Feb. 2018 - Dec. 2018, 344 observations for stocktwits; Feb-Jul. 2018, 181 observations for reddit), respectively. Exogenous variables include message volume, return volatility and weekday dummies. p-values are in parentheses. The feedback of R_t on the lagged S_t is particularly reported. For a sake of brevity, the estimates from its own lagged term are suppressed.

	Bubble-growing period		post-Bubble period	
S_t	StockTwits	Reddit	StockTwits	Reddit
R_{t-1}	0.1313	0.094	0.2294	0.2288
	(0.000)	(0.000)	(0.000)	(0.024)
R_{t-2}	0.0885	0.063	0.1306	-0.0053
	(0.007)	(0.003)	(0.026)	(0.957)
R_{t-3}	0.0153	0.021	0.1187	0.044
	(0.6388)	(1.015)	(0.029)	(0.643)
adjustR^2	0.944	0.927	0.690	0.603

Table 10: Feedback effects from return to sentiment with market microstructure effects

This table documents a mutual feedback system constituted by the CRIX return (R_t) and sentiment measure (S_t), for the bubble-growing period (Aug. 2014 - Jan. 2018, 1280 observations) and for the post-bubble period (Feb. 2018 - Dec. 2018, 344 observations for stocktwits; Feb-Jul. 2018, 181 observations for reddits), respectively. Exogenous variables include message volume, return volatility and weekday dummies. p-values are in parentheses. The feedback of S_t on the lagged R_t is particularly reported. For a sake of brevity, the estimates from its own lagged term are suppressed.

6.2 Out-of-sample Forecasts

Existing literature has shown that variables that perform well in-sample do not necessarily perform well out-of-sample (Spiegel, 2008). Ultimately, out-of-sample tests are more relevant for assessing the genuine predictive power of forecasters and are much less affected by the econometric issues such as the over-fitting concern. In this section, we are particularly interested to compare the performance of sentiment predictability between the bubble and non-bubble period. Garcia (2013) find that the predictability of stock returns using news content is concentrated in recessions, which seems consistent with the literature from psychology and economics that investors sensitivity to news is most pronounced when they are going through hard times (Tetlock et al. (2008)). Their results suggest that investors' opinion is more likely to be affected by news sentiment during recessions hence moving the market. Different from news articles, our online message data directly captures investors' opinions. In this case, the tests of sentiment predictability (based on the online messages) on future market return conditional on the bubble and post-bubble period answers the question when the investors' opinion is more consistent with investors' trading behavior. Before conducting the analysis, we expect that during the bubble period, investors should be more likely to trade on sentiment, given the fact that the trading volume is usually higher in the bubble period than the non-bubble period. We then investigate the out-of-sample predictive performance of the CL-based sentiment to answer the question.

Following Spiegel (2008), the out-of-sample forecast of next periods expected CRIX return is recursively computed as:

$$\hat{R}_{t+1} = \hat{\alpha}_t + \hat{\beta}_t S^{CL,StockTwits}_{1:t;t} \tag{8}$$

where $\hat{\alpha}_t$ and $\hat{\beta}_t$ are the ordinary least squares estimates by regressing $\{R_{t+1}\}_{t=1}^{T-1}$ on a constant and the CL-based sentiment, $\{S^{CL,StockTwits}_{1:t;t}\}_{t=1}^{T-1}$, with the initial estimation window from 2014-08-01 to 2014-12-31. The in-sample forecast is computed as same as above except that $\hat{\alpha}_t$ and $\hat{\beta}_t$ are replaced by those estimated by using the entire sample. Like our in-sample analogues, we consider alternative sentiment measures for comparison purposes and report the results in Table 11.

To evaluate the out-of-sample forecasting performance, we apply the pervasive Campbell and Thompson (2007) R^2_{OS} statistics based on the unconstrained forecast and truncated forecast im-

Predictor	R^2_{OS}	CW-test	DM-test
Whole Sample	0.6571	3.0665***	2.0982**
Bubble Period	0.9703	3.4135**	2.4929***
Post-bubble Period	8.7603	3.1022***	2.2031**

Table 11: Out-of-sample Forecasting

This table reports the out-of-sample performances of CL-based sentiment in predicting the daily excess cryptocurrency market return for the whole sample, bubble and non-bubble period respectively. The preditive regression is estimated recursively using the data available at the forecast formation time t. R^2_{OS} is the out-of-sample R^2 following Campbell and Thompson (2007). CW-test is the Clark and West (2007) MSFE-adjusted statistic calculated according to prevailing mean model. DM-test is the modified Diebold and Mariano (2002) t-statistic. The initial estimation window starts from 2014-08-01 to 2014-12-31 for the whole sample and bubble period while it starts from 2018-01-01 to 2018-05-31 for the post-bubble period. The out-of-sample period starts from 2015-01-01 to 2018-12-27, 2015-01-01 to 2017-12-31 and 2018-06-01 to 2018-12-27 for the whole sample, bubble period and post-bubble period respectively. *, **, and *** indicate significance at the 10%, 5% and 1% levels respectively.

posing non-negative equity premium constraint. The unconstrained R^2_{OS} statistic measures the proportional reduction in Mean Squared Forecast Error (MSFE) for the predictive regression forecast relative to the historical average benchmark. Welch and Goyal (2007) shows that the historical average is a very stringent out-of-sample benchmark, and individual economic variables typically fail to outperform the historical average. To compute R^2_{OS}, let s be a fixed number chosen for the initial sample training, so that the future expected return can be estimated at time $t = s+1, s+2, ..., T$. Then, we compute $T-s$ out-of-sample forecasts: $\{\hat{R}_{t+1}\}_{t=s}^{T-1}$. Note that we use the data over 2014-08-01 to 2014-12-31 as the initial estimation window so that the out-of-sample period spans over 2015-01-01 to 2018-12-27.

$$\hat{R}^2_{OS} = 1 - \frac{\sum_{t=r}^{T-1}(R_{t+1} - \hat{R}_{t+1})^2}{\sum_{t=r}^{T-1}(R_{t+1} - \bar{R}_{t+1})^2}, \tag{9}$$

where $\bar{R}_{t+1}$ denotes the historical average benchmark corresponding to the constant expected return model ($R_{t+1} = \alpha + \epsilon_{t+1}$), i.e.,

$$\bar{R}_{t+1} = \frac{1}{t}\sum_{s=1}^{t} R_s. \tag{10}$$

By construction, the R^2_{OS} statistic lies in the range $(-\infty, 1]$. If $R^2_{OS} > 0$, it means that the forecast $\hat{R}_{t+1}$ outperforms the historical average $\bar{R}_{t+1}$ in terms of MSFE.

The statistical significance of the out-of-sample R^2s we report is based on MSFE-adjusted statistic of Clark and West (2007) (CW-test hereafter) and Diebold and Mariano, 2002 statistic (DM-test hereafter). The null hypothesis under the CW test is that the historical average

MSFE is not greater than the predictive regression forecast MSFE. The test is formulated as $\mathcal{H}_0 : R^2_{OS} \leq 0$ against $\mathcal{H}_1 : R^2_{OS} > 0$. Clark and West (2007) shows that the test has a standard normal limiting distribution when comparing forecasts from the nested models. Comparing a predictive regression forecast to the historical average is a kind of nested model, as the predictive regression model reduces to the historical average under the null hypothesis. Under the null hypothesis that the constant expected return model generates a series, the predictive regression model produces a noisier forecast than the historical average benchmark as it estimates the slope parameters with zero population values. We thus expect the benchmark models MSFE to be smaller than the predictive regression model's MSFE under the null. The MSFE-adjusted statistic accounts for the negative expected difference between the historical average MSFE and predictive regression MSFE under the null so that it can reject the null even if the R^2_{OS} statistic is negative.

Table 11 shows that overall our sentiment measure serves as a strong predictor for the CRIX return with a whole sample R^2_{OS} of 0.66. Surprisingly, the results show that sentiment predictor seems to predict future market return much stronger in a non-bubble period than the bubble period. For example, compared to the benchmark model, the CL-based sentiment generates a significant positive R^2_{OS} of 8.76% for the post-bubble period while it is only 0.97% for the bubble period. It implies that during a post-bubble period, incremental predictability relative to the simple average strategy is contributed to the sentiment measure. The fact that the benchmark can perform during the bubble period can be linked to the positive feedback effort of De Long et al. (1990b).

6.3 Trading strategies

In this section, we evaluate the economic significance of a trading strategy based on the predictability of CL-based sentiment conditional on the bubble and post-bubble periods. The intuition is straightforward as a success of sentiment in forecasting returns suggests investors who utilize our crypto-specific lexicon can develop profitable trading strategies based on daily variation in the Cryptocurrency market. Following Pesaran and Timmermann (1995), Cochrane et al. (2007), Rapach et al. (2010) and Detzel et al. (2018), we define the buy indicator (buy=1) of sentiment strategy as:

$$\textbf{Buy}_{type,source,t} = \begin{cases} 1, & \text{if out-of-sample forecasted return} > 0 \\ 0, & \text{Otherwise} \end{cases} \tag{11}$$

The out-of-sample forecasted return is estimated recursively based on the predictive regression $R_{t+1} = \alpha + \sum_{i=0}^{i=p-1} \beta_i S_{t-i}^{type,source} + e_t$, which only includes investor sentiment as predictors. In particular, we include up to 7 lags of investor sentiment to fully capture the return reversal effects. The return on the CRIX sentiment strategy on day t is given by:

$$R_t^{type,source} = Buy_{type,source,t} \times R_t + (1 - Buy_{type,source,t}) \times R_{f,t} \tag{12}$$

where R_t and $R_{f,t}$ denote the CRIX return and the risk-free rate at time t respectively. Table 12 presents summary statistics for the buy-and-hold and sentiment strategies. Overall, the results suggest that CL-based sentiment provides valuable return predictability to generate better investment value comparing to an equal weight trading strategy and the buy-and-hold trading strategy. For example, using the whole sample, the CL-based sentiment strategy appears to achieve the highest portfolio returns (28bps daily return), which is more than twice of equal weight portfolio return and 1.5 times of buy-and-hold CRIX return. Consistent with out-of-sample tests, the CL-sentiment performs well in the post-bubble period in terms of CRIX return. Although in the bubble period, the CL-sentiment achieves 37bps daily portfolio return, it is not surprising given the CRIX return is 40bps. On the contrary, during the post-bubble period, the CRIX return is -45bps while CL-based sentiment maintains a positive investment value as a sharp comparison. In this case, we believe that our crypto-specific lexicon indeed gains value to the portfolio managers during the market downturn. In the meantime, it further confirms our argument that sentiment plays an important role in predicting future Cryptocurrency market returns. The superiority of sentiment strategy is especially highlighted given that the benchmark momentum strategy fades out during the market downturn.

	Mean	SD	Sharp	Min	Max	Skew	Kurt	MDD
Panel A: Whole Sample								
$S^{CL,StockTwits}$	0.0028	0.0330	0.0846	-0.2533	0.1985	-0.6660	8.8141	0.7353
EW	0.0010	0.0199	0.0480	-0.1266	0.0993	-0.7834	6.1501	0.6818
CRIX	0.0019	0.0398	0.0474	-0.2533	0.1985	-0.784	6.1505	0.9195
Panel B: Bubble Period								
$S^{CL,StockTwits}$	0.0037	0.0296	0.1234	-0.2384	0.1985	-0.3422	11.2475	0.4202
EW	0.0020	0.0183	0.1092	-0.1192	0.0993	-0.6039	7.6384	0.2668
CRIX	0.0040	0.0366	0.1088	-0.2384	0.1985	-0.6045	7.6397	0.4868
Panel C: Post-bubble Period								
$S^{CL,StockTwits}$	0.0002	0.0414	0.0039	-0.2533	0.1216	-0.8825	4.6221	0.7353
EW	-0.0022	0.0239	-0.0927	-0.1266	0.0657	-0.8449	3.2899	0.6818
CRIX	-0.0045	0.0477	-0.0940	-0.2533	0.1313	-0.8449	3.2898	0.9195

Table 12: Horse race between alternative predictors

This table presents summary statistics of the returns of Sentiment Cryptocurrency portfolio strategies in excess of the 1-day risk-free rate. The trading strategy is formed based on a long position in CRIX if out-of-sample forecasted CRIX return ≥ 0, and the risk-free rate otherwise. EW denotes an equal-weighted portfolio between CRIX and risk-free interest rate. The sample period is daily from 2014-08-01 - 2018-12-27. The initial estimation window is from 2014-08-01 - 2014-12-31 for the whole sample period and bubble period while it is 2018-01-01 - 2018-05-31 for the post-bubble period. Mean, SD (standard deviation), SR (sharpe ratio), Min (minimum value), Max (maxiumum value), Skew (skewness), Kurt (kurtosis) and MDD (maximum drawdown) are reported for each trading strategy.

7 Conclusion

Analyzing a very large dataset of several millions of messages discussed on social media (StockTwits) and message boards (Reddit), we provide empirical evidence showing that specific attention had to be paid when measuring investor sentiment for a new type of asset such as cryptocurrencies, and that the impact of investor sentiment on cryptocurrency returns is conditional on bubble regimes.

First, investor sentiment helps to forecast future cryptocurrency returns only when investor sentiment is quantified using the appropriate methodology and a domain-specific lexicon. In that regard, we show that using a crypto-specific lexicon to take into account the specificity of the language used by unsophisticated investors on the crypto market is of a utmost importance, as many new terms, such as blockchain, ICO, hackers, wallet, shitcoin, binance, and hodl, are not covered in existing financial or psychological dictionaries. Emojis, such as 🚀 (positive) 💩 (negative), also play a very important role in capturing the exact sentiment of messages published on social media.

Second, investor sentiment helps to forecast future cryptocurrency returns mostly when investor sentiment is derived from a dataset of messages related to the financial aspect of cryptocurrency. Investor sentiment can be distilled using many sources of textual content, such as articles from traditional media, user-generated content on message boards, or short-messages on social media. We provide evidence that, at least in a market driven by individual investors who possess higher risk preference, and when there is only limited information about the fundamental value of the underlying asset, measuring investor sentiment from messages related to the financial aspect of cryptocurrency gives better results than considering discussions about the technology (blockchain, mining, wallet) and the economic implication of cryptocurrencies. This finding confirms that what drives the attention of online investors is mostly the evolution of prices and not the evolution of the technology.

Last, we provide empirical evidence showing that, contrary to the stock market, investor sentiment in the crypto market is not reversed over the next few weeks. Using a local-momentum autoregression model, we find that the sentiment effect is prolonged and sustained during the bubble, while the sentiment effect turns to a reversal effect once the bubble collapsed. During the bubble, optimistic sentiment causes a price soar, and a subsequent positive-feedback or momentum trading given price soars inflates investor sentiment further. Interestingly, a trading strategy based on investor sentiment further shows that sentiment is particularly important for the portfolio managers to get rid of the market downturn during the post-bubble period. Overall, our results might have some important implication to better understand how bubbles form following investors' illusion for new technologies.

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.

Athey, S., Parashkevov, I., Sarukkai, V., and Xia, J. (2016). Bitcoin pricing, adoption, and usage: Theory and evidence.

Avery, C. N., Chevalier, J. A., and Zeckhauser, R. J. (2016). The CAPS prediction system and stock market returns. *Review of Finance*, 20(4):1363–1381.

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

Balduzzi, P., Das, S. R., and Foresi, S. (1998). The central tendency: A second factor in bond yields. *Review of Economics and Statistics*, 80(1):62–72.

Balvers, R. J. and McDonald, B. (2017). Designing a global digital currency.

Böhme, R., Christin, N., Edelman, B., and Moore, T. (2015). Bitcoin: Economics, technology, and governance. *Journal of Economic Perspectives*, 29(2):213–38.

Brody, S. and Diakopoulos, N. (2011). Cooooooooooooooollllllllllllll!!!!!!!!!!!!!!: Using word lengthening to detect sentiment in microblogs. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '11, pages 562–570, Stroudsburg, PA, USA. Association for Computational Linguistics.

Bukovina, J., Marticek, M., et al. (2016). Sentiment and Bitcoin volatility. Technical report, Mendel University in Brno, Faculty of Business and Economics.

Campbell, J. Y. and Thompson, S. B. (2007). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531.

Cheah, E.-T. and Fry, J. (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, 130:32–36.

Chen, C. Y.-H., Härdle, W. K., Hou, A. J., and Wang, W. (2018). Pricing cryptocurrency options: the case of crix and bitcoin. *revise and resubmitted to Journal of Financial Econometrics.*

Chordia, T., Roll, R., and Subrahmanyam, A. (2001). Market liquidity and trading activity. *The Journal of Finance*, 56(2):501–530.

Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311.

Cochrane, J. H., Longstaff, F. A., and Santa-Clara, P. (2007). Two trees. *The Review of Financial Studies*, 21(1):347–385.

Cooper, M. J., Dimitrov, O., and Rau, P. R. (2001). A rose. com by any other name. *The Journal of Finance*, 56(6):2371–2388.

da Gama Silva, P. V. J., Klotzle, M. C., Pinto, A. C. F., and Gomes, L. L. (2019). Herding behavior and contagion in the cryptocurrency market. *Journal of Behavioral and Experimental Finance*, 22:41–50.

Das, S. R. and Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science*, 53(9):1375–1388.

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

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990b). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance*, 45(2):379–395.

Deng, S., Sinha, A. P., and Zhao, H. (2017). Adapting sentiment lexicons to domain-specific social media texts. *Decision Support Systems*, 94:65 – 76.

Detzel, A. L., Liu, H., Strauss, J., Zhou, G., and Zhu, Y. (2018). Bitcoin: Learning, predictability and profitability via technical analysis.

Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 20(1):134–144.

Duan, J.-C. (2016). Local-momentum autoregression and the modeling of interest rate term structure. *Journal of Econometrics*, 194(2):349–359.

Easley, D., O'Hara, M., and Basu, S. (2017). From mining to markets: The evolution of bitcoin transaction fees.

Feldman, R., Govindaraj, S., Livnat, J., and Segal, B. (2010). Managements tone change, post earnings announcement drift and accruals. *Review of Accounting Studies*, 15(4):915–953.

Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance*, 68(3):1267–1300.

Geuder, J., Kinateder, H., and Wagner, N. F. (2018). Cryptocurrencies as financial bubbles: The case of Bitcoin. *Finance Research Letters*.

Hafner, M. C. (2018). Testing for bubbles in cryptocurrencies with time-varying volatility. *Journal of Financial Econometrics*.

Harvey, C. R. (2014). Bitcoin myths and facts.

Jermann, U. J. (2018). Bitcoin and cagan's model of hyperinflation.

Kim, S.-H. and Kim, D. (2014). Investor sentiment from internet message postings and the predictability of stock returns. *Journal of Economic Behavior & Organization*, 107:708 – 729. Empirical Behavioral Finance.

Lerman, A. and Livnat, J. (2010). The new form 8-k disclosures. *Review of Accounting Studies*, 15(4):752–778.

Liu, Y. and Tsyvinski, A. (2018). Risks and returns of cryptocurrency. Technical report, National Bureau of Economic Research.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1):35–65.

Loughran, T. and McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230.

Oliveira, N., Cortez, P., and Areal, N. (2016). Stock market sentiment lexicon acquisition using microblogging data and statistical measures. *Decision Support Systems*, 85:62 – 73.

Pagnotta, E. and Buraschi, A. (2018). An equilibrium valuation of bitcoin and decentralized network assets.

Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up?: Sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10*, EMNLP '02, pages 79–86, Stroudsburg, PA, USA. Association for Computational Linguistics.

Peng, L. (2005). Learning with information capacity constraints. *Journal of Financial and Quantitative Analysis*, 40(2):307–329.

Pesaran, M. H. and Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. *The Journal of Finance*, 50(4):1201–1228.

Rapach, D. E., Strauss, J. K., and Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies*, 23(2):821–862.

Renault, T. (2017). Intraday online investor sentiment and return patterns in the u.s. stock market. *Journal of Banking & Finance*, 84:25 – 40.

Spiegel, M. (2008). Forecasting the equity premium: Where we stand today. *The Review of Financial Studies*, 21(4):1453–1454.

Sprenger, T. O., Tumasjan, A., Sandner, P. G., and Welpe, I. M. (2014). Tweets and trades: the information content of stock microblogs. *European Financial Management*, 20(5):926–957.

Sun, L., Najand, M., and Shen, J. (2016). Stock return predictability and investor sentiment: A high-frequency perspective. *Journal of Banking & Finance*, 73:147–164.

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.

Tetlock, P. C., Saar-Tsechansky, M., and Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, 63(3):1437–1467.

Trimborn, S. and Härdle, W. K. (2018). CRIX an index for cryptocurrencies. *Journal of Empirical Finance*, 49:107–122.

Welch, I. and Goyal, A. (2007). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4):1455–1508.

Yermack, D. (2017). Corporate governance and blockchains. *Review of Finance*, 21(1):7–31.