Association for Information Systems

AIS Electronic Library (AISeL)

ICIS 2019 Proceedings

DLT, Blockchain and FinTech

Trading on Cryptocurrency Markets: Analyzing the Behavior of Bitcoin Investors

Alexander Keller
University of Passau, precisionconference@kelleralexander.de

Michael Scholz
University of Passau, michael.scholz@uni-passau.de

Trading on Cryptocurrency Markets: Analyzing the Behavior of Bitcoin Investors

Completed Research Paper

Alexander Keller
University of Passau
Passau, Germany
alexander.keller@uni-passau.de

Michael Scholz
University of Passau
Passau, Germany
michael.scholz@uni-passau.de

Abstract

Driven by innovative information technologies, the financial industry is facing a recent disruptive fintech revolution. One emerging technology within this field is cryptocurrency, aiming to change the future means of payment. In this paper, we study Bitcoin exchange trading and examine what factors influence the behavior of different cryptocurrency investor types. To answer this question, market bids are considered in form of investors' offers and orders as a proxy for their trading behavior. First, an unsupervised clustering technique is applied in order to group different types of investors based on similarities in trading behavior. Second, a supervised classification mechanism is used on social media news to measure the sentiment influencing trading decisions. Among other indicators this bullishness is integrated in an autoregressive distributed lag (ARDL) model to identify the factors influencing the trading behavior of investor types. Besides large investors, foreign traders and speculators, cryptocurrency-specific market participants are characterized in the form of miners. With identifying indicators driving investors' actions (i.e., macro-financial fundamentals, technical trading indicators, technological measures and market sentiment), this study contributes to recent research by explaining the trading behavior on cryptocurrency markets and its impact on exchange rates.

Keywords: Financial analytics, cryptocurrency, fintech markets, trading behavior, trading strategies, investor types, tweet bullishness, SIMEX

Introduction

Cryptocurrencies like Bitcoin, Litecoin or Ethereum have received great public attention due to the exceptional volatility of exchange rates. Supporters see cryptocurrencies as the future means of payment with some organizations like Expedia and Microsoft already starting to accept them. In contrast, others give warning and see the hype as a driver for a speculative bubble (Brezo and Bringas 2012; Garcia et al. 2014).

Cryptocurrencies are digital assets managed by a network, which in general utilizes distributed ledger technology and encryption to log, control and verify both transactions and creation of new currency units, i.e., coins or tokens. Like physical currencies, e.g., the US-Dollar (USD), the digital coins are primarily used as a medium of exchange. However, cryptocurrencies are neither backed by any central bank or government institution nor do they rely on any economic value creation processes. In comparison to other assets like equity or bonds, returns from cryptocurrency investments can only be generated via a rise in price. Therefore, the digital asset is strongly driven by speculations which are reflected in a high volatility of the exchange rate (Glaser et al. 2014; Kristoufek 2013).

Besides investors, the new technology also raises attention from researchers, who investigate the underlying blockchain (e.g., Risius and Spohrer 2017; Underwood 2016), legal issues of cryptocurrencies (e.g., Bryans 2014; Pieters and Vivanco 2017) and economic and financial aspects (e.g., Blau 2018; Kristoufek 2015; Mai et al. 2018). Prior studies primarily conducted econometric analyses using price-related dependent variables (e.g., Ciaian et al. 2016; Li and Wang 2017). However, with the Wall Street Journal reporting that individual investors dominate cryptocurrency markets in contrast to other financial markets (Osipovich 2018), price determinants of traditional assets might not be valid in this field. Besides market fundamentals and technological indicators, scholars especially focus on analyzing the effect of social media on the price of cryptocurrencies. Although evidence is provided that social media news significantly impact the price of Bitcoin (Garcia and Schweitzer 2015), the relationship between social media sentiment and price fluctuations is seen as complex (Mai et al. 2018).

Additionally, recent research on cryptocurrencies has considered investors as homogeneous in terms of their trading behavior, which is contrary to the finding of disparate types of investors reacting differently to given information (Li et al. 2018). Therefore, we argue that the observed effects of social media as well as other determinants on cryptocurrency exchange rates depend on those investor types.

With this study, we contribute to research by addressing three key issues: (i) Which types of investors are trading cryptocurrencies and how do they differ? (ii) What indicators are influencing the trading behavior of these investor types? (iii) Which types of investors drive the cryptocurrency exchange rate?

To answer the first question, we use an unsupervised clustering technique to identify types of investors based on information derived from market bids on a cryptocurrency exchange platform. Covering the second question, we study the impact of market sentiment in form of bullishness as well as macro-financial and cryptocurrency-related indicators on the trading behavior of identified investor types. We address the third question by analyzing the influence of different investor types on the exchange rate of a cryptocurrency. While the quantitative data is collected from Kraken[1], Yahoo! Finance[2] and Blockchain.info[3], qualitative information about the market sentiment is assessed via the application programming interface of Twitter[4].

As results, we observe ten different types of investors trading Bitcoins. Due to the market bids' characteristics, we identify trading types like cryptocurrency miners, large investors and speculators on the supply side. While speculators show a herd behavior driven by bullishness and the number of tweets, large investors' trading is based on technical indicators like the relative strength index and macro-financial fundamentals like the Crude Oil price and the USD index. Additionally, we show that the behavior of only some types of investors is influencing the exchange rate of Bitcoin.

Related Work

As cryptocurrencies, such as Bitcoin, are traded on financial markets, we first review findings about the role of information in general on financial markets. Thereafter, we review the role of investors' mood on their trading behavior. Finally, we review literature that investigates determinants of cryptocurrency markets (i.e., volume and price) that might also determine the behavior of cryptocurrency traders.

The Role of Information in Financial Markets

In 1970 the efficient-market hypothesis (EMH) was introduced by Eugene Fama, who states that the price of a security fully reflects all information available in the market (Fama 1970). The theory argues that market actors have complete access to information, and hence new insights instantly lead to a trading behavior which results in a price adjustment on the market. Due to this market efficiency, it is impossible to consistently outperform a market as prices always incorporate all relevant information available.

---

[1] *Kraken* represents the biggest cryptocurrency exchange platform with a registered location in the United States.

[2] *Yahoo! Finance* is a news site that provides financial data and reports.

[3] *Blockchain.info* is a block explorer service that provides data about the Bitcoin blockchain.

[4] *Twitter* is a leading microblogging service. Microblogs shared via the platform are called *tweets*.

Starting in the 1980s, the consistency of this theory was discussed in academic research and empirical evidence was provided which describes market anomalies that contradict EMH (Brown et al. 1983; Cohen et al. 2003; Grossman and Stiglitz 1980). On the one side predictable patterns were identified based on valuation and asset-specific parameters. For example, Campbell and Shiller (1988) statistically test dividend yields to predict stock prices and show that long-horizon stock returns are forecastable.

Based on the econometric analyses of quantitative parameters, the field of behavioral finance emerged in the 1990s, considering models of human psychology (Shiller 2003). Within this field, De Long et al. (1990) distinguish between two different individual investor types: arbitrageurs and noise traders. Arbitrageurs are informed individuals who drive prices with a close relationship to underlying fundamentals. Noise traders, in contrast, react on pseudo-signals like sentiment and popular trading strategies, deciding irrationally and erratic (Black 1986).

A behavioral effect in form of feedback has been studied in market environments, where speculative prices increase within a short time period, which “attracts public attention, promotes word-of-mouth enthusiasm, and heightens expectations for further price increases” (Shiller 2003). This feedback-trading can result in price momentum-effects for single stocks (Jegadeesh and Titman 1993) as well as indices (Chan et al. 2000). With the ease of internet trading, one can see an increase of uninformed individual investors, who adopt these feedback trading strategies (Xiaoquan and Lihong 2015). This phenomenon also serves as an explanation for the emergence of financial bubbles, which are strongly driven by market speculations that result in feedback-cycles (Marimon et al. 1993; Smith et al. 1988).

Along with the emergence of stock message boards, the impact of information from virtual communities on trading behavior was studied by Park et al. (2013). As some investors are observed to prefer messages that support their prior beliefs, Park et al. (2013) reveal a confirmation bias and conclude that participation in stock message boards does not necessarily lead to higher returns. The concept of disagreement is related to the confirmation bias and represents another factor influencing investors’ behavior. Although disagreement about expected returns can induce trading in terms of volume (Harris and Raviv 1993; Kandel and Pearson 1995; Karpoff 1986), it can also result in the no-trade theorem (Milgrom and Stokey 1982). In this situation individuals reconsider their price and market opinion because of disagreement instead of trading with one another.

The Impact of Investor Mood

Recent studies in the field of behavioral finance show that investment decisions are strongly affected by emotional impulses of individuals based on their mood, also known as sentiment (Bollen et al. 2011; De Long et al. 1990; Li et al. 2014). A widely used representation for investors’ mood are textual information sources like news reports (Feuerriegel and Prendinger 2016; Schumaker et al. 2012; Tetlock 2007) or message boards (Antweiler and Frank 2004; Das and Chen 2007). Wysocki (1998) was among the first who considered stock message postings to forecast stock market volatility. He predicts the next-day trading volume based on message postings from Yahoo! Finance. Antweiler and Frank (2004) use message postings to calculate a bullishness measure derived from computational linguistic methods. They found a significant influence of messages’ bullishness on stock volatility. Das and Chen (2007) are the first to propose a methodology based on classification algorithms to systematically extract investor opinion from textual information sources. They give empirical evidence by generating a sentiment index for 24 tech sector stocks and found relation between their index and market activity.

With the emergence of social media platforms, the sentiment of microblog messages was considered to predict financial market movements. Bollen et al. (2011) can significantly improve the accuracy in predicting market changes of the Dow Jones Industrial Average (DJIA) by including specific public mood dimensions from Twitter and Google. Oliveira et al. (2016) demonstrate that tweet sentiment and volume are relevant to forecast returns of the S&P 500. Li et al. (2014) predict the stock movement on firm-level and propose a technique to mine Twitter for sentiment analysis. Sprenger et al. (2014) also find an association between tweet sentiment and returns, and state that negative information in tweets has a stronger impact on stock prices than positive information. In addition, they emphasize that tweets serve as “valuable proxies for investor behavior and belief formation” (Sprenger et al. 2014). Meta-information about tweets (e.g., amount of followers) is included by Nofer and Hinz (2015), who develop a trading strategy for the German stock market and increase their portfolio value by approximately 36% within half a year. Besides Twitter, other social media platforms have been investigated. Karabulut (2013) uses Facebook's Gross National Happiness to predict changes in daily returns and trading volume in the US stock market. Chen et al. (2014) study user-generated content from Seeking Alpha, a content service platform for financial markets. Their results show that future stock returns can be predicted with articles and reader commentaries from the platform. In addition, Ho et al. (2017) show that the relationship between social media sentiment and stock returns is time-varying.

Determinants of Cryptocurrency Markets

Recently, cryptocurrency exchanges have drawn attention from researchers who are interested in determinants that influence the cryptocurrency market in terms of price, volatility, and transaction volume. Within this field, influencing factors fall into three categories: (i) indicators about the macro-financial market condition, (ii) indicators reflecting cryptocurrency-related fundamentals and technological aspects, and (iii) indicators of cryptocurrency market sentiment which serve as a proxy for market attractiveness.

The first category includes general economic measures (e.g., interest rates, oil price) as a proxy for the macro-economic condition and financial market measures (e.g., DJIA, S&P 500) as indicators for financial market development. Li and Wang (2017) demonstrate that since 2014 the long-term Bitcoin exchange rate is more sensitive to macro-financial indicators than to technological indicators. This is in line with the results from Sovbetov (2018), who identifies a negative short-term correlation between Bitcoin prices and the S&P500 index. In contrast, Ciaian et al. (2016) cannot detect any long-term effects of macro-financial indicators like the oil price or the DJIA on the Bitcoin price. Existing literature suggests that most cryptocurrency-related studies use a set of those fundamentals as independent variables to control for macro-financial condition and development (e.g., Blau 2018; Kristoufek 2015; Mai et al. 2018).

The second category includes cryptocurrency-related fundamentals (e.g., trading volume, volatility) and technological measures (e.g., hash rate[5]). Although there is no evidence that volume can predict the volatility of returns (Balcilar et al. 2017), trading volume, as well as volatility, are identified as significant determinants of exchange rates in both short- and long-term perspective (Sovbetov 2018). While the hash rate has a significant influence, the Bitcoin transaction volume is not affecting the price (Bouoiyour and Selmi 2015). The results are supported by Kristoufek (2015), who observes a positive correlation of the hash rate on the Bitcoin price in the long-term. However, the relationship is unstable and becomes weaker over time. Li and Wang (2017) explore mining difficulty, which is correlated with the hash rate, and detect a negative long-term effect of difficulty and Bitcoin price without considering time trends. They also show that the price primarily responds to trading volume and exchange rate volatility.

The third category covers market sentiment indicators to measure the perceived investment attractiveness of a cryptocurrency market. While measures like Wikipedia views or Google search queries are used as a proxy of investors' attention (Kristoufek 2013), Mai et al. (2018) apply bullishness as a combined measure of the quantity of positive and negative forum postings to consider the investment attractiveness. Based on messages from BitcoinTalk.org, they show that more bullish posts are associated with higher future Bitcoin returns and trace this effect back to the silent majority in the social network. As a result, they state that social media sentiment can predict the future price of Bitcoin, but mention that the relationship is complex. This is supported by Garcia and Schweitzer (2015), who include social signals based on Twitter tweets in a Bitcoin trading algorithm and confirm social media sentiments' potential for positive trading returns.

Theoretical Model

In this paper, we argue that the trading behavior of cryptocurrency investors is, in general, heterogeneous. However, homogeneous types can be identified, who react similarly to particular indicators. As a proxy for

[5] Measuring unit for the available computing power within the Bitcoin network [terahash/sec].

behavior, we consider market bids in terms of offers and orders. With this, we assume that investors actual market bids are a more precise gauge for the behavior of investors than the exchange rate. This arises from the mechanism that calculates the exchange rate based on settled transactions, i.e., an offer or order has been accepted by a buyer or seller, while market bids directly represent investors' actions.

Built upon prior research (see in Section "Related Work"), information influencing trading behavior can be classified into three dimensions: The first category includes indicators in form of economic and financial information as a proxy for the general macro-financial condition. The second category consists of cryptocurrency-related information, in terms of technical as well as trading indicators as a proxy for cryptocurrency market activity. The bullishness indicator and tweet count represent a proxy for the market sentiment, which describe the third category.

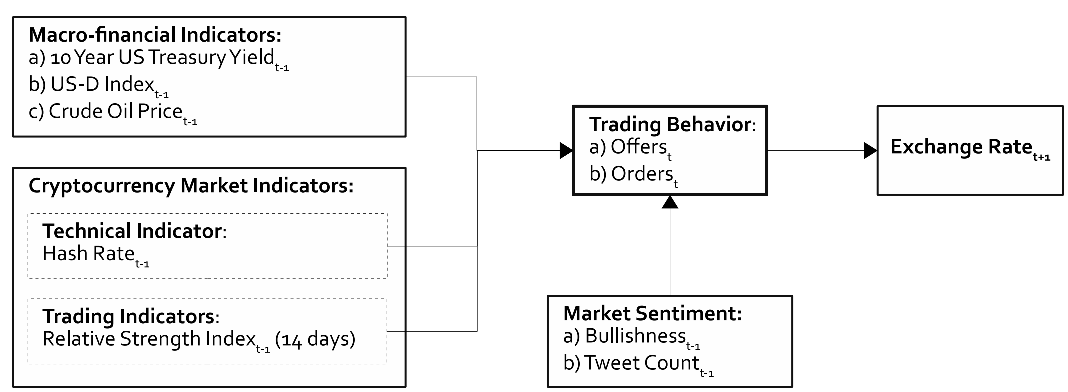

Figure 1. Theoretical Model

As different types of investors respond differently to information sources (Li et al. 2018), our theoretical model (see Figure 1) investigates the influence of information and investors' mood on the behavior of investors orders and offers. The orders and offers represent the trading behavior. These translate into market transactions, and thus directly determine the exchange rate.

Empirical Data

This study is based on four different types of information: (i) text-based microblog messages derived from news accounts on Twitter as a representation of market mood towards cryptocurrencies, (ii) trading data in the form of marked bids and completed transactions from the cryptocurrency exchange, Kraken, (iii) general finance data representing the macro-financial condition extracted from Yahoo! Finance, and (iv) Technical information about the underlying blockchain technology gathered via Blockchain.info. For the purpose of the latter analysis, data was collected between February and May of 2018, a total timeframe of 101 days.

For the collection of microblog messages, we focused on 25 well-known financial news accounts on Twitter, which publish in English language and have a large audience according to their follower count. We choose Twitter as data source because it represents the platform with the broadest acceptance in the financial community (Li et al. 2018). We collected data from financial news accounts, rather than streams in, order to ensure (i) a high information quality, (ii) a high number of potential readers including potential traders that are likely to transform the messages into action, and (iii) a high credibility. With nine of 25 accounts having a special focus on cryptocurrency-related content, we cover a wide range of financial news and aim to ensure the representation of most publicly available information about cryptocurrencies. All considered news sites have at least 20,000 followers. In total, 234,758 tweets were collected, whereby 24,547 contained cryptocurrency-related news. Within this subset, daily news tweets range from 1 to 225, averaging to 25.18 messages per day with a standard deviation of 38.26. For accessing the data, we use the Twitter application programming interface (API).

The trading data was gathered from the cryptocurrency exchange, Kraken. We collected information about market bids in terms of offers and orders as well as trades via the platforms public market data API. Market bids are specified via a unique identifier (id), a timestamp, volume and price. Additionally, we collected the duration of a market bid to gather the amount of time a bid was available on the market. A trade represents a market transaction and besides the attributes id, timestamp, volume and price, it is described by a specific type (i.e., buy or sell) and mode (i.e., market or limit order). Within the time period of 101 Bitcoin trading days, we observed 2.77 million trades, 4.69 million offers and 4.38 million orders. The volume weighted average exchange rate was 8,792.77 USD per Bitcoin with a standard deviation of 1,368.68 USD. Tukey's five for the daily price and volume of all successful trades, all offers, and all orders are presented in Table 1.

Table 1. Cryptocurrency Exchange Data (Descriptive Statistics)

	Amount	Daily Price (in USD)					Daily Volume (in BTC)				
		Min	Q25	Q50	Q75	Max	Min	Q25	Q50	Q75	Max
Trades	2,771,495	6,651	8,119	8,739	9,433	11,528	2,209	4,4084	7,321	11,023	22,753
Offers	4,685,035	6,612	8,056	8,708	9,357	11,449	78,704	107,724	123,896	145,340	531,433
Orders	4,377,868	6,711	8,157	8,846	9,507	11,595	84,468	111,291	124,935	138,435	197,461

We obtained daily financial data as a proxy for the macro-financial condition from Yahoo! Finance. While the 10-year US Treasury yield represents the attractiveness of an alternative investment, the USD Index acts as a strength indicator for the exchange rate currency in which Bitcoin is traded. Besides US-specific indicators, we use the Oil price as proxy for the global macro-economic condition. When we observe missing values due to weekends and bank holidays, we replace these by a weighted moving average (Moritz et al. 2015). We apply a weighting where factors decrease exponentially and use a semi-adaptive window size of four to ensure that the replacement is based on eight observations (i.e., four values to the left and right of the missing value).

Technical information about the underlying Bitcoin blockchain is gathered via Blochchain.info. As the proof-of-work difficulty only changes every 14 days for the Bitcoin blockchain, we use the daily hash rate as a proxy for the activity in the Bitcoin network.

Data Preparation and Analysis

For the purpose of the econometric analysis, we prepare the collected data and conduct some pre-analysis steps in order to generate dependent variables as input for the econometric models.

Tweet Bullishness Classification

The data collected from Twitter is used to derive an indicator of the market sentiment. For the preparation of the gathered textual data, we take several steps in order to prepare the textual data for classification of supposed trading signals provided in tweets.

In the first step, we reduce the messages by removing components which make the information readable for humans and components which are not necessary for evaluating the meaning of a tweet. (a) At the beginning, retweets and identical tweets are identified and removed. (b) All remaining messages are transformed into lower case and punctuation, apostrophes and newlines are eliminated. (c) Finally, URLs linked to the tweets are omitted.

In the second step, we reduce the tweets' feature space by tokenizing repeating patterns in messages. Besides the usage of hyperlinks, tweets on Twitter can be linked to other users and to content-related topics. In the first case (a) the username is prefixed with the sign '@' which indicates that the message is related to a specific account. In the second case (b) words are prefixed with the sign '#' or '$' representing a linkage to a certain topic. While '#' is used in general, '$' is used in financial contexts, indicating a linkage to a ticker symbol (e.g., $BTC stands for Bitcoin). Both signs are replaced by an equivalent token. (c) Besides the identification of Twitter-related components, negation is tokenized to negate the meaning of tweets (e.g., 'doesn't' is changed to 'does xxxnegationxxx'). (d) Finally, we replace emoticons with a positive or negative token (e.g., 'bought some coins today :-)' with 'bought some coins today xxxpositivexxx'). We additionally test stemming to remove morphological endings from words and delete English stop words, but do not observe an increase in accuracy by applying these techniques to prepare the textual data.

In order to predict the sentiment signals from the gathered news tweets, we randomly draw a subset of cryptocurrency-related tweets[6] and manually classify the messages as buy, hold or sell recommendations. Coding is done by the two authors and a third coder with cryptocurrency trading experience. In order to ensure the robustness of training data, we only consider tweets that are classified unanimously by all three coders. In total, the training set includes 1000 tweets of which 63.7% are hold recommendations. Among the remainder, sell signals are more likely (20.3%) than buy recommendations (16%). Due to this relation, the training set is unbalanced and consists of a large number of variables (i.e., each word identified in the tweets is used as a variable in the classification). Therefore, we apply the nearest shrunken centroids classifier which is introduced by Tibshirani et al. (2003) to solve high-dimensional classification problems. Based on a ten-fold cross validation, an accuracy of over 71% is achieved (in sample accuracy over 90%).

Based on the training, we predict the trading signal of 24,547 gathered cryptocurrency-related news tweets. In order to use these recommendations in the econometric analysis, we aggregate the buy and sell signals in a bullishness index called $TweetBullishness_t$ with t as time interval (based on Antweiler and Frank 2004; Li et al. 2018). M_t^c represents the number of messages, with $c \in \{buy, sell\}$. The measure is defined as:

$$TweetBullishness_t = log\left(\frac{1 + M_t^{buy}}{1 + M_t^{sell}}\right) \tag{1}$$

The Bullishness index reflects the number of buy recommendations in relation to the number of sell signals. While a negative value indicates a time interval with more sell signals, positive values imply a higher amount of buy recommendations.

Clustering Investor Types

The gathered data about a market bid's volume, price, time and duration provides information about the trading behavior of market participants. In order to answer the first research question about different types of cryptocurrency investors, we utilize the market bid information and characterize investor types on the supply (i.e., market offers) and demand side (i.e., market orders). With applying an unsupervised clustering technique, we aim to group market bids based on similarities. As the price-related input variable, we use a market bid's relative price (i.e., bid price divided by exchange rate) at the time of bid placement. We assume that if an investor wants to sell his assets quickly, he or she places an offer severely under the market price and vice versa. The second variable considered is a market bid's volume and serves as a proxy for the financial power of a trader. Finally, we include the duration of a market bid in the clustering. This variable represents the amount of time a market bid is available on the exchange and stands for the timing of a market bid relative to the market condition.

---

[6] These news tweets include words related to cryptocurrencies (i.e., crypto, coin, bitcoin, etc.).

As a clustering technique, we apply a model-based approach based on finite mixture modelling in which 14 different Gaussian models are estimated with an expectation maximization algorithm (Fraley and Raftery 2002). The number of clusters and the covariance parameterization is selected on the Bayesian Information Criterion (BIC). As parameters, we specify the minimum (3) and maximum amount (7) of clusters for which the BIC is calculated. Additionally, we initialize the clustering with a subset of market bids. The results are presented in the next section.

Econometric Model

We propose that several indicators directly affect the behavior of cryptocurrency investors rather than the exchange rate. The exchange rate (i.e., price) of a cryptocurrency might be driven by only some types of investors. Following this argumentation, we first estimate the effect of several indicator variables on the number of placed bids at a particular day from each identified cluster. In a second analysis, we investigate the effect of each investor cluster on the exchange rate of Bitcoin.

Our first model analyzes the relationship between investment indicators and investors' behavior. We use the cluster-day panel data generated with the previously described data preparation and pre-analysis steps.

This analysis is conducted with the following model:

$$\Delta BidCount_{t,c} = \beta_1 \Delta BidCount_{t-1,c} + \beta_2 Weekend_t + \Gamma X_{t-1} + \lambda Week_t + \epsilon_{t,c} \quad (2)$$

The dependent variable $\Delta BidCount_{t,c}$ is the first difference of the number of placed bids by investor cluster c at day t regarding the day before. We use an autoregressive distributed lag (ARDL) model and assume that the data generating process is nested in the ARDL(1,1) model (see evidence for lag structure in Section "Model Diagnostics"). A high volatility of the Bitcoin exchange rate provides a first indication for using such a narrow lag structure to capture dynamic effects. Variable $Weekend$ captures weekend effects and is set to 1 if the actual day is either a Saturday or a Sunday and otherwise it is set to zero. The vector X_{t-1} represents all investment indicators described in Section "Theoretical Model". Their effects are captured in vector Γ. We further integrate time-fixed effects per week.

The vector of indicator factors consists of variables whose values are objectively determined (e.g., the USD index), and one variable (tweet bullishness) whose values are inevitably estimated with some error. Thus, the results of our econometric models for the daily bid count differences might be biased when not correcting for this estimation error. Yang et al. (2018) propose two simulation-based methods for error correction, SIMEX and MC-SIMEX. Either the error variance (SIMEX) or the misclassification error of a classification (MC-SIMEX) are needed to run these simulation-based methods. The tweet bullishness aggregates the bullishness of all tweets published on a particular day (see Equation 1). Tweet bullishness, thus, is not the result of a single classification, but a measure derived from a classification combined with an aggregation. Hence, MC-SIMEX is not suitable in this context. We instead apply SIMEX as an error correction method to the first econometric model (see Equation 2). We compute the error variance based on a simulation with 10,000 iterations. Therefore, we first compute the confusion matrix for the tweet classification. This confusion matrix holds the number of randomly generated messages that are considered as true positives, false positives, true negatives and false negatives. In each simulation iteration, we generate a certain number of messages that definitely indicate buying Bitcoins and a certain number of messages that definitely indicate selling Bitcoins. The number of bullish and bearish type messages is determined by the distribution of each respective type found in the manually tagged tweets. Both distributions are of type Weibull ($shape_{Sell} = 1.5170$, $scale_{Sell} = 10.3268$, $shape_{Buy} = 1.2721$, $scale_{Buy} = 7.4710$). The true bullishness score is computed based on Equation 1, and the number of definitely bullish as well as the number of definitely bearish messages. The confusion matrix determines how many of the messages per simulation iteration are classified as buy, sell and hold messages. The observed bullishness score is then estimated based on the simulated message classifications. Therefore, the tweet bullishness error per simulation iteration is determined by the difference between the true and the observed bullishness. The error variance, finally, is the variance across the error values of all simulation iterations.

In order to get robust regression coefficients, we run SIMEX with 1,000 bootstrap replications and the computed error variance of 0.8006 for tweet bullishness per regression model (i.e., per cluster).

Our second model analyzes the relationship between the investors' behavior and the exchange rate between Bitcoin and USD. We again operationalize each investor type's behavior by the number of placed bids per day. The dependent variable is the daily exchange rate difference (i.e., close rate open rate). We estimate the following model with time-fixed effects per week and weekend effects:

$$\Delta Rate_t = \beta_1 Weekend_t + \beta_2 Close_{t-1} + \Gamma Z_{t-1} + \lambda Week_t + \epsilon_t \quad (3)$$

The vector Z_{t-1} includes the cluster specific logarithm of the number of placed bids on day $t-1$. The effect of each investor type on the exchange rate is hence captured in vector Γ. We include the close price at day $t-1$ as an additional covariate.

Results

Investor Types

We identified six investor types that place offers and four investor types that place orders. An investor might belong to multiple investor types, because these types are identified based on behavioral patterns rather than personal characteristics. All identified types of investors placing offers are shown in Table 2.

Table 2. Types of Investors that Placed Offers

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Proportion	0.1792	0.3270	0.2902	0.0557	0.1337	0.0142
Average Bid Volume (in BTC)	2.4503	0.3687	3.3483	6.9861	5.8943	15.8013
Average Relative Price	0.9969	0.9971	0.9897	0.9781	0.9881	0.9487
Average Duration (in min)	0.0000	0.0000	0.2683	38.7902	3.4647	3100.8057

Most offers belong to Cluster 1, 2 or 3. These clusters describe investors that offer a rather small number of Bitcoins at a rather high price. Cluster 1 places offers with a medium-low volume at approximately the current market price. Cluster 2 is characterized by offers with a very small volume and an average duration of 0, indicating that this cluster represents Bitcoin miners that aim at selling their Bitcoins at the highest possible price in order to cover their mining costs. Cluster 3 deviates from clusters 1 and 2 by a higher bid volume and a lower offer price. Clusters 4 and 5 represent investors that place medium-sized offers. The major difference between these two clusters is the duration of an offer listed on *Kraken* before it was successfully executed. Investors in either of these clusters also set a significantly lower price than investors that belong to one of the first three clusters. Cluster 6 covers investors that offer a rather large number of Bitcoins. They consequently charge a lower price and must wait more than 2 days on average before their offers are successfully accepted. Cluster 6 thus seems to represent large professional investors.

All identified investor types on the demand side are described in Table 3. The average relative price of all order types is larger than 1 indicating that all investors are willing to pay more per Bitcoin than the actual market price. Traders in Cluster 7 are willing to pay the highest price per Bitcoin. However, orders within this cluster are online for more than one day. This indicates that investors of Cluster 7 place their orders at a time where there are no corresponding investors that want to sell their Bitcoins. Clusters 8 and 10 cover investors that want to buy a midsize volume of Bitcoins. Their major difference is the price they are willing to pay per Bitcoin. Cluster 8 places orders at a price that is quite close to the current exchange rate. Their orders are hence successfully executed within a minute. Most investors belong to Cluster 9 which covers investors that want to buy a rather small number of Bitcoins. Cluster 10 is the smallest cluster on the demand side.

Table 3. Types of Investors that Placed Orders				
	Cluster 7	Cluster 8	Cluster 9	Cluster 10
Proportion	0.0234	0.3333	0.5437	0.0995
Average Bid Volume (in BTC)	3.1531	5.3156	1.0808	5.1934
Average Relative Price	1.0411	1.0092	1.0031	1.0174
Average Duration (in min)	1898.8899	0.6159	0.0000	12.1059

Compared to the supply side, there is no corresponding investor type that is trading large amounts of Bitcoins. This indicates that the Bitcoin market was rather bearish during our observation period.

Indicator Impact on Investors' Behavior

We analyzed the effect of several indicators on the first difference of placed bids per cluster with our first econometric model (Equation 2) and the SIMEX approach. The results for the six investor types that place order bids are presented in Table 4. The adjusted R^2 values and the information criteria were calculated based on residuals, because log-likelihood values are not available with SIMEX (Yang et al. 2018).

Table 4. Indicators Influencing Investor Types Offering Bitcoins						
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
ΔBidCount$_{t-1}$	-0.383***	-0.275**	-0.208***	-0.114***	-0.367***	-0.166***
Weekend$_t$	-1025.858***	-1001.167***	-1206.337**	-280.663	-526.638**	-74.899
TweetBullishness$_{t-1}$	-603.500*	293.759	-625.703	164.686	276.364	55.431
TweetCount$_{t-1}$	-33.194*	7.985	54.327	-13.750	-5.710	-5.581
RSI(14)$_{t-1}$	-9.811	-65.900	-40.960	42.897	-9.422	14.518*
USDollar$_{t-1}$	503.980	1555.740**	3192.708***	-1075.967*	666.398**	-217.240*
USTreasury$_{t-1}$	-5195.815	-6019.683	-19956.359**	-12589.111**	-5574.799	-2016.492
Oil$_{t-1}$	269.771	168.136	978.018***	-1244.815***	149.848	-217.240*
BitcoinHash$_{t-1}$	-195.694	896.073	-842.537	-898.260*	-312.653	-243.037**
Adj. R^2	*0.106*	*0.332*	*0.440*	*0.830*	*0.132*	*0.748*
AIC	*1710.148*	*1814.489*	*1865.902*	*1733.046*	*1674.680*	*1440.137*
BIC	*1775.773*	*1880.114*	*1931.526*	*1798.670*	*1740.304*	*1505.762*

The behavior of each investor type on the supply side is affected by at least one indicator. Trading decisions of investors in Cluster 1 largely depend on market mood (i.e., tweet bullishness and tweet count) as well as the number of bids placed on the day before. These traders might be speculators showing a herd behavior that is strongly susceptible to market mood transported via social media channels. Traders of type 2 belong to the largest cluster on the supply side and are characterized by a very small average volume of placed offers. Investors that belong to this cluster mainly consider the USD index to decide whether to offer Bitcoins. The higher the USD index, the more coins are offered by investors that belong to Cluster 2 compared to the day before. A higher USD index is especially worthwhile for investors who sell Bitcoins in USD but need to exchange the gained dollars in another currency (e.g., the Euro). Considering the low average bid count, the large cluster size and the influence of the USD, we characterize these traders as global cryptocurrency miners bringing their Bitcoins to market. Investors of type 3 can be seen as informed traders strongly relying on macro-financial indicators. Besides the oil price, the USD index is positively influencing the daily bid count difference.

As these measures provide an early indication of inflationary development, they signal potential changes in the general price level affecting Bitcoin trading. Hence, a higher value of the USD relative to some foreign currencies induces more offers placed by global investors. As we observe a positive effect for Cluster 3, we assume that this cluster represents not only informed but also global traders. In contrast, Cluster 4 shows the opposite effect, indicating USD orientated investors. Additionally, investors of Cluster 4 place bids regarding the number of Bitcoins generated per day, indicating informed traders offering medium volumes of Bitcoins (average bid volume 6.9861). Like Cluster 3, investors of Cluster 5 place more offers when the USD is valued higher relative to foreign currencies. This implicates that traders benefit from a high exchange rate when changing USD into their preferred currency. However, the average bid volume from Cluster 5 is significantly larger than the one from Cluster 3. Additionally, it seems that traders of Cluster 5 are less attracted to alternative investments. Investors of Cluster 6 are well-informed market actors and the only ones to sell more Bitcoins if the relative strength index (RSI) over 14 days indicates that the market is overbought, and Bitcoins should be offered rather than ordered. These investors also place more offers when the Bitcoin hash rate is low and thus, the number of generated Bitcoins per day is lower. As indicated by a negative USD index and oil price coefficient, Cluster 6 covers USD-orientated investors trading the largest bid volumes on average.

Table 5 shows the model results for the four investor types that order Bitcoins. In contrast to the behavior of investor types on the supply side, the behavior of investors ordering Bitcoins only sparsely depends on our proposed indicators.

Table 5. Indicators Influencing Investor Types Ordering Bitcoins

	Cluster 7	Cluster 8	Cluster 9	Cluster 10
$\Delta BidCount_{t-1}$	-0.338**	-0.130	-0.136	-0.410***
$Weekend_t$	-54.410	-1105.544*	-1987.343***	-227.930
$TweetBullishness_{t-1}$	-23.735	644.711	91.856	100.678
$TweetCount_{t-1}$	-2.914	23.201	-35.655	0.630
$RSI(14)_{t-1}$	0.120	-44.318	63.147	-5.334
$USDollar_{t-1}$	14.328	2370.052**	-263.638	327.374
$USTreasury_{t-1}$	-1498.748	-8986.486	-13663.502	-3480.397
Oil_{t-1}	-39.498	516.399	-736.547	-257.344***
$BitcoinHash_{t-1}$	-76.690	291.734	1839.524	-65.047
Adj. R^2	*-0.081*	*0.236*	*0.214*	*0.242*
AIC	*1366.005*	*1880.490*	*1910.827*	*1611.836*
BIC	*1431.629*	*1946.114*	*1976.452*	*1677.460*

Investors in Cluster 7 do not seem to consider macro-financial indicators, cryptocurrency market indicators or market sentiment. This does not mean that these investors are uninformed, but that they either consider other indicators or do not rely on any market information. In contrast to Clusters 8 and 9, investors of type 7 do not place significantly more orders during weekdays than on weekends. Cluster 8 is characterized by a very high positive coefficient for the USD index indicating that traders in this cluster rather focus on USD investments. Additionally, their trading activities are significantly lower on weekends. Cluster 9 represents more than 50% of orders and shows the best market timing (i.e., mean duration of zero) in trading. Although we cannot observe any influence of indicators, these investors appear to be largely professional as indicated by a significant weekend effect. Investors of Cluster 10 place significantly more orders when the oil price decreases. With an average bid volume of over five Bitcoins, they might use the oil price as an indicator for the global market condition. As seen in Cluster 7, these investors are negatively influenced by the difference of bid counts of the day before, which results in a non-persistent trading behavior.

Impact of Investor Types on the Exchange Rate

As shown in the previous sections, the behavior of most identified investor types is driven by at least one of the considered indicators. With our second econometric model (see Equation 3), we aim to identify those investor types who drive the Bitcoin-USD exchange rate. Table 6 shows that the daily exchange rate difference is mainly and significantly driven by the number of orders placed by Cluster 7. We furthermore found a weakly significant effect on the daily exchange rate difference for the number of orders published by investors of Cluster 8.

Table 6. Influence of Investor Types on the Daily Exchange Rate Difference

	Estimate	Std. Error	t-Value	p-Value
$Close_{t-1}$	-4169.023	859.314	-4.852	<0.001***
$Weekend_t$	87.991	76.964	1.143	0.257
$BidCount_{t-1,1}$	15.306	351.258	0.044	0.965
$BidCount_{t-1,2}$	-168.031	482.909	-0.348	0.729
$BidCount_{t-1,3}$	-72.326	304.713	-0.237	0.813
$BidCount_{t-1,4}$	456.967	372.374	1.227	0.224
$BidCount_{t-1,5}$	-35.389	437.383	-0.081	0.936
$BidCount_{t-1,6}$	9.946	238.889	0.042	0.967
$BidCount_{t-1,7}$	-854.943	283.889	-3.012	0.004**
$BidCount_{t-1,8}$	-650.847	330.376	-1.970	0.053
$BidCount_{t-1,9}$	-152.665	534.886	-0.285	0.776
$BidCount_{t-1,10}$	118.265	370.353	0.319	0.750
Adj. R^2	*0.260*			
AIC	*1476.660*			
BIC	*1549.883*			

In summary, we found the Bitcoin exchange rate to be driven by orders rather than offers. However, the investor types influencing the daily exchange rate are those who do not consider any indicators when deciding when to place a bid.

Model Diagnostics

Our dependent variables (cluster-specific differences of placed bids on two subsequent days and exchange rate on two subsequent days) are generated by a stationary process as indicated by ADF-GLS and KPSS tests. The ADF-GLS test is an augmented Dickey-Fuller test applied to non-trending data without intercept. The null hypothesis is tested that a unit root is present in the data. The KPSS test, in contrast, tests the null hypothesis that the data is drawn from a stationary process, and hence, no unit root is present. The test statistics shown in Table 7 clearly indicate stationarity of all our dependent variables.

Autocorrelation function values as well as partial autocorrelation function values indicate that the memory of all cluster-specific daily bid count differences is one whereas there is no memory for the daily exchange rate difference. We thus specified the lag structure as presented in Equations 2 and 3.

Table 7. Model Diagnostics						
	ADF-GLS Test		KPSS Test		Box-Pierce Test	
	Statistic	p-Value	Statistic	p-Value	Statistic	p-Value
$\Delta BidCount_1$	-9.0351	<0.001	0.2581	>0.1	1.7965	>0.1
$\Delta BidCount_2$	-7.3987	<0.001	0.2999	>0.1	1.8981	>0.1
$\Delta BidCount_3$	-7.3042	<0.001	0.3512	>0.1	0.4232	>0.1
$\Delta BidCount_4$	-10.4022	<0.001	0.0324	>0.1	1.6388	>0.1
$\Delta BidCount_5$	-5.9023	<0.001	0.0921	>0.1	0.5867	>0.1
$\Delta BidCount_6$	-9.0040	<0.001	0.1296	>0.1	0.0680	>0.1
$\Delta BidCount_7$	-5.5512	<0.001	0.0517	>0.1	1.2828	>0.1
$\Delta BidCount_8$	-9.6110	<0.001	0.3199	>0.1	0.0042	>0.1
$\Delta BidCount_9$	-9.1146	<0.001	0.2251	>0.1	0.1549	>0.1
$\Delta BidCount_{10}$	-8.1943	<0.001	0.0421	>0.1	1.5907	>0.1
$\Delta ExchangeRate$	-7.0681	<0.001	0.1022	>0.1	0.3368	>0.1

We estimated ARDL models for each dependent variable. Endogeneity is unlikely to be an issue in an ARDL model as soon as the error terms are serially uncorrelated. The Box-Pierce test demonstrates that we can assume uncorrelated error terms (see Table 7). The null hypothesis of a Box-Pierce test is that the data are serially uncorrelated. Variance inflation factors less than five for each model and each independent variable furthermore indicate the absence of multi-collinearity in our econometric models.

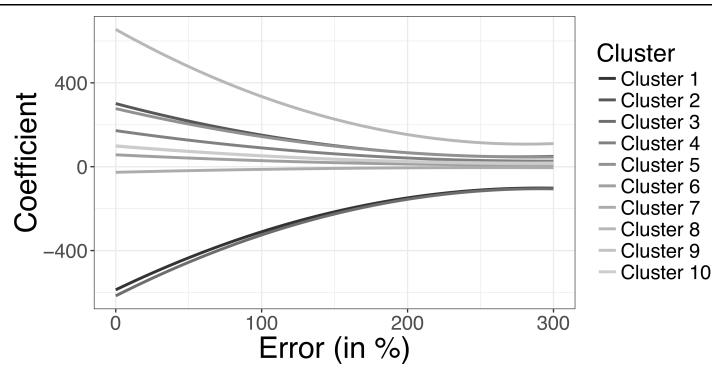

Figure 2. Impact of Classification Errors on the Estimated Coefficients

We used SIMEX to correct for errors in the variable tweet bullishness that stem from misclassifications. Figure 2 shows the impact of different classification error levels on the estimated coefficient for tweet bullishness. The measured classification error is represented by a tweet bullishness error of 100% in Figure 2. A classification approach that will produce more classification errors than the applied nearest shrunken neighbor centroid approach would produce a tweet bullishness error of more than 100%. SIMEX presumably reduces the tweet bullishness error to 0% (Yang et al. 2018). The coefficients for tweet bullishness at an error of 0% thus match the coefficients as estimated with SIMEX. Figure 2 shows that the tweet bullishness coefficient is largely affected by SIMEX in some models (i.e., for Clusters 1, 3, and 8). We thus can assume that the applied error correction approach – SIMEX – was effective.

Discussion

Our study provides evidence that the behavior of Bitcoin traders is influenced by several macro-financial, cryptocurrency-related and market sentiment indicators. However, bitcoin traders are not sensitive to these indicators in the same manner. In total, we identified ten different trading types, six types of investors offering Bitcoins and four types of investors ordering Bitcoins. Investors are especially influenced by macro-financial indicators, such as the USD index or the oil price when placing offers. Investors' decisions to place orders seem to be only sparsely influenced by market mood, macro-financial and technical indicators. It is this rather vague dependence on indicators and, particularly, investor types who place orders that drive Bitcoin's exchange rate. This causes the Bitcoin market to lack transparency and invokes investors to take speculative approaches.

These findings clearly indicate the Bitcoin traders' behavior is significantly different from the behavior of traders in other markets. The major difference is that the behavior of Bitcoin traders is rather arbitrary, leading to a high volatility of the Bitcoin exchange rate. This makes it more difficult for researchers and practitioners to analyze the characteristics of the Bitcoin market and derive trading strategies, as well as regulation policies. Our study gives some interesting insights into the characteristics of the market and provides interesting findings for researchers investigating the Bitcoin market, traders that want to build an efficient trading strategy and regulation institutions. More specifically, the knowledge about the investor types helps i) explain events, such as significant exchange rate changes, ii) evaluate efficiency and stability of the Bitcoin market, iii) develop trading strategies, and iv) predict future exchange rates.

Theoretical Implications

With this study, we contribute to research by providing several theoretical implications. First, we show that Bitcoin trading is strongly driven by the demand side, and support the findings from Ciaian et al. (2016). We additionally show that investors' decisions to place offers are largely influenced by investment indicators, whereas investors on the demand tend to decide without considering these indicators. Bitcoin trading thus is driven by erratic investor buying decisions. Second, by presenting speculators (i.e., Cluster 1) influenced by market mood and large professional investors (i.e., Cluster 6) driven by fundamentals, we provide evidence, that the distinction between arbitrageurs and noise traders from De Long et al. (1990) is also valid in Fintech markets. However, we identified further investor types in the form of e.g., miners (i.e., Cluster 2) and rather uninformed global investors (i.e., Cluster 5), who play an important role on the supply side of cryptocurrency markets. Third, our results confirm the effect of social media founded in related research (e.g., Antweiler and Frank 2004; Mai et al. 2018). But, with showing that only some types of investors are driven by market mood indicators, we supplement prior findings and are able to explain the influence in more detail. More specifically, our results show that market mood indicators, such as the tweet bullishness, do not necessarily have an impact on the price of a cryptocurrency. In our dataset, we found that only one type of investor (i.e., Cluster 1) is affected by the market mood indicator. However, these investors do not affect the Bitcoin price. By using SIMEX when integrating the bullishness indicator into the econometric models, we additionally guarantee the robustness of the provided results. Based on SIMEX, we found that a decrease of the tweet bullishness by one unit will increase the daily difference of the bid count of Cluster 1 by 603 bids. Without SIMEX, we would have estimated that one unit decrease in the tweet bullishness would lead to an increase of only 316 bids. Taking the results by Yang et al. (2018) into account, we assume that the estimates based on SIMEX better approximate the true coefficients. Finally, we show that the trading behavior of different types of investors can be explained with the examined indicators. We hence argue that this behavior is deterministic rather than stochastic.

Practical Implications

Cryptocurrencies have become an important issue not only for investors but also for financial institutions and regulatory authorities. Our results provide several implications for investors and regulators and help to better understand the dynamics in the Bitcoin market. On the one hand, one can derive certain trading strategies. For example, when investors want to buy a rather moderate or large volume of Bitcoins, they have a fairly high chance to find adequate offers at a rather low price if the oil price as well as the USD index decreases. Vice versa, investors offering Bitcoin should consider the return of alternative investments like US treasuries as well as the first difference of placed offers on the day before. Additionally, large volumes should be split into smaller parts, in order to ensure a quick trading of market bids. On the other hand, the results can support decision making as to whether cryptocurrencies should be regulated due to security issues and speculative trading concerns. The findings help to grasp the behavior of different market actors and therefore generate a general understanding of the market functionalities. Additionally, the study provides evidence that some types of investors (i.e., Cluster 1) are showing a herd behavior which might result in speculative market movements. By identifying only one rather small order cluster with a significant influence on the exchange rate, we did not find evidence that a specific trading type can influence price movements.

Limitations and Future Research

Our research is subject to three limitations. First, as Bitcoin price data differs across exchanges (Pieters and Vivanco 2017), the observed results are limited to trading activities on Kraken. Although the exchange is one of the largest, future research should investigate other trading platforms to discuss our results. Secondly, we only gathered data from Bitcoin investors who actively placed bids. Investors who merely passively react on market bids and accept an order or an offer are not represented within our dataset. These investors do not leave any publicly available data on trading platforms, making it hard to investigate factors influencing their behavior. Although it is unclear which indicators these passive investors rely on, it seems that their actions do not drive the exchange rate. This is because their reactions always depend on placed market bids, i.e., the studied behavior of investors. And third, we considered offers and orders to be independent of each other in our study. Several investors might, however, place multiple bids and the timing might also depend on other bids. As offers and orders are made publicly available without any information that allows disclosing the identity of an investor, additional non-anonymized data is required to overcome this issue and get further insights into the behavior of Bitcoin traders. Although we provided evidence that the studied indicators do have an influence on the trading behavior, we did not analyze the interdependence between the behavior and the exchange rate in detail. Further research could focus on this aspect and explain the effects on the exchange rate.

References

Antweiler, W. and Frank, M. Z. 2004. “Is all that talk just noise? The information content of Internet stock message boards,” Journal of Finance (59:3), pp. 1259–1294.

Bagnoli, M., Beneish, M. D., and Watts, S. G. 1999. “Whisper Forecasts of Quarterly Earnings Per Share,” Journal of Accounting and Economics (28:1), pp. 27–50.

Balcilar, M., Bouri, E., Gupta, R., and Roubaud, D. 2017. “Can volume predict Bitcoin returns and volatility? A quantiles-based approach,” Economic Modelling (64), pp. 74–81.

Black, F. 1986. “Noise,” The Journal of Finance (41:3), p. 529.

Blau, B. M. 2018. “Price dynamics and speculative trading in Bitcoin,” Research in International Business and Finance (43), pp. 15–21.

Bollen, J., Mao, H., and Zeng, X. 2011. “Twitter mood predicts the stock market,” Journal of Computational Science (2:1), pp. 1–8.

Bouoiyour, J. and Selmi, R. 2015. “What does Bitcoin look like?,” Annals of Economics and Finance (16:2), pp. 449–492.

Brezo, F. and Bringas, P. G. 2012. “Issues and Risks Associated with Cryptocurrencies such as Bitcoin,” in Proceedings of the Second International Conference on Social Eco-Informatics, pp. 20–26.

Brown, P., Kein, D. B., Kleidon, A. W., and Marsh, T. A. 1983. “Stock return seasonalities and the tax-loss-selling hypothesis: Analysis of the arguments and Australian evidence,” Journal of Financial Economics (12:1), pp. 105–127.

Bryans, D. 2014. “Bitcoin and Money Laundering: Mining for an Effective Solution Bitcoin and Money Laundering: Mining for an Effective Solution,” Indiana Law Journal (89:1), pp. 441–472.

Campbell, J. Y. and Shiller, R. J. 1988. “Stock Prices, Earnings, and Expected Dividends,” The Journal of Finance (43:3), pp. 661–676.
Chan, K., Hameed, A., and Tong, W. 2000. “Profitability of Momentum Strategies in the International Equity Markets,” The Journal of Financial and Quantitative Analysis (35:2), p. 153.
Chen, H., De, P., Hu, Y., and Hwang, B. H. 2014. “Wisdom of crowds: The value of stock opinions transmitted through social media,” Review of Financial Studies (27:5), pp. 1367–1403.
Ciaian, P., Rajcaniova, M., and Kancs, D. 2016. “The economics of BitCoin price formation,” Applied Economics (48:19), pp. 1799–1815.
Cohen, R. B., Polk, C., and Vuolteenaho, T. 2003. “The Value Spread,” Journal of Finance (58:2), pp. 609–641.
Das, S. R. and Chen, M. Y. 2007. “Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web,” Management Science (53:9), pp. 1375–1388.
De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. 1990. “Noise Trader Risk in Financial Markets,” Journal of Political Economy (98:4), pp. 703–738.
Fama, E. F. 1970. “Efficient Capital Markets-A Review of Theory and Empirical Work,” The Journal of Finance (25:2), pp. 383–417.
Feuerriegel, S. and Prendinger, H. 2016. “News-based trading strategies,” Decision Support Systems (90), pp. 65–74.
Forgas, J. P. 1995. “Mood and Judgment: The Affect Infusion model (AIM),” Psychological Bulletin (117:1), pp. 39–66.
Fraley, C. and Raftery, A. E. 2002. “Model-Based Clustering, Discriminant Analysis, and Density Estimation,” Journal of the American Statistical Association (97:458), pp. 611–631.
Garcia, D. and Schweitzer, F. 2015. “Social Signals and Algorithmic Trading of Bitcoin,” Royal Society Open Science (2), pp. 1–13.
Garcia, D., Tessone, C. J., Mavrodiev, P., and 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).
Glaser, F., Zimmermann, K., Haferkorn, M., Weber, M. C., and Siering, M. 2014. “Bitcoin - Asset or Currency? Revealing Users’ Hidden Intentions,” in Proceedings of the Twenty Second European Conference on Information Systems,
Grossman, S. J. and Stiglitz, J. E. 1980. “On the Impossibility of Informationally Efficient Markets,” American Economic Review (70:3), pp. 393–408.
Harris, M. and Raviv, A. 1993. “Differences of Opinion Make a Horse Race,” The Review of Financial Studies (6:3), pp. 473–506.
Hirshleifer, D. 2001. “Investor Psychology and Asset Pricing,” The Journal of Finance (56:4), pp. 1533–1597.
Ho, C. S., Damien, P., Gu, B., and Konana, P. 2017. “The time-varying nature of social media sentiments in modeling stock returns,” Decision Support Systems (101), pp. 69–81.
Jaffe, J. F., Westerfield, R., and Ma, C. 1989. “A twist on the Monday effect in stock prices. Evidence from the U.S. and foreign stock markets,” Journal of Banking and Finance (13:4-5), pp. 641–650.
Jegadeesh, N. and Titman, S. 1993. “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency,” The Journal of Finance (48:1), pp. 65–91.
Kandel, E. and Pearson, N. D. 1995. “Differential Interpretation of Public Signals and Trade in Speculative Markets,” Journal of Political Economy (103:4), pp. 831–872.
Karabulut, Y. 2013. “Can Facebook Predict Stock Market Activity?,” Working Paper (October 21, 2013).
Karpoff, J. M. 1986. “A Theory of Trading Volume,” The Journal of Finance (41:5), pp. 1069–1087.
Keim, D. B. and Stambaugh, R. F. 1986. “Predicting returns in the stock and bond markets,” Journal of Financial Eco- nomics (17:2), pp. 357–390.
Kristoufek, L. 2013. “BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era,” Scientific Reports (3:3415), pp. 1–7.
Kristoufek, L. 2015. “What are the main drivers of the bitcoin price? Evidence from wavelet coherence analysis,” PLoS ONE (10:4), pp. 1–15
Li, B., Chan, K. C. C., Ou, C., and Ruifeng, S. 2017. “Discovering public sentiment in social media for predicting stock movement of publicly listed companies,” Information Systems (69), pp. 81–92.

Li, Q., Wang, T., Li, P., Liu, L., Gong, Q., and Chen, Y. 2014. "The effect of news and public mood on stock movements," Information Sciences (278), pp. 826–840.

Li, T., Van Dalen, J., and Van Rees, P. J. 2018. "More than just noise? Examining the information content of stock microblogs on financial markets," Journal of Information Technology (33:1), pp. 50–69.

Li, X. and Wang, C. A. 2017. "The technology and economic determinants of cryptocurrency exchange rates: The case of Bitcoin," Decision Support Systems (95), pp. 49–60.

Mai, F., Shan, Z., Bai, Q., Wang, X., and 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), pp. 19–52.

Marimon, R., Spear, S. E., and Sunder, S. 1993. "Expectationally driven market volatility: An experimental study," Journal of Economic Theory (61:1), pp. 74–103.

Milgrom, P. and Stokey, N. 1982. "Information, trade and common knowledge," Journal of Economic Theory (26:1), pp. 17–27.

Moritz, S., Sardá, A., Bartz-Beielstein, T., Zaefferer, M., and Stork, J. 2015. "Comparison of different Methods for Uni- variate Time Series Imputation in R," CoRR (abs/1510.03924).

Nofer, M. and Hinz, O. 2015. "Using Twitter to Predict the Stock Market: Where is the Mood Effect?," Business and Information Systems Engineering (57:4), pp. 229–242.

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

Osipovich, A. 2018. "Little Guys and Big Trading Firms Square Off in Bitcoin Futures Arena," in The Wall Street Journal.

Park, J. H., Konana, P., Gu, B., Kumar, A., and Raghunathan, R. 2013. "Information valuation and confirmation bias in virtual communities: Evidence from stock message boards," Information Systems Research (24:4), pp. 1050–1067.

Pieters, R. G. and Vivanco, S. 2017. "Financial regulations and price inconsistencies across Bitcoin markets," Information Economics and Policy (39), pp. 1–14.

Risius, M. and Spohrer, K. 2017. "A Blockchain Research Framework," Business & Information Systems Engineering (59:6), pp. 385–409.

Schumaker, R. P., Zhang, Y., Huang, C.-N., and Chen, H. 2012. "Evaluating sentiment in financial news articles," Decision Support Systems (53), pp. 458–464.

Shiller, R. J. 2003. "From Efficient Markets Theory to Behavioral Finance," Journal of Economic Perspectives (17:1), pp. 83–104.

Smith, V. L., Suchanek, G. L., and Williams, A. W. 1988. "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets," Econometrica (56:5), pp. 1119–1151.

Sovbetov, Y. 2018. "Factors Influencing Cryptocurrency Prices: Evidence from Bitcoin, Ethereum, Dash, Litcoin, and Monero Journal of Economics and Financial Analysis," Journal of Economics and Financial Analysis (2:2), pp. 1– 27.

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), pp. 926–957.

Tetlock, P. C. 2007. "Giving content to investor sentiment: The role of media in the stock market," Journal of Finance (62:3), pp. 1139–1168.

Thaler, R. 1985. "Mental Accounting and Consumer Choice," Marketing Science (4:3), pp. 199–214.

Thaler, R. H. 1987. "Anomalies: The January Effect," Journal of Economic Perspectives (1:1), pp. 197–201.

Tibshirani, R., Hastie, T., Narasimhan, B., and Chu, G. 2003. "Class Prediction by Nearest Shrunken Centroids, with Applications to DNA Microarrays," Statistical Science (18:1), pp. 104–117.

Underwood, S. 2016. "Blockchain beyond bitcoin," Communications of the ACM (59:11), pp. 15–17.

Wysocki, P. D. 1998. "Cheap Talk on the Web: The Determinants of Postings on Stock Message Boards," Working Paper (November 1998).

Xiaoquan, Z. and Lihong, Z. 2015. "How Does the Internet Affect the Financial Market? an Equilibrium Model of Internet- Facilitated Feedback Trading," MIS Quarterly (39:1), 17–A10.

Yang, M., Adomavicius, G., Burtch, G., and Ren, Y. 2018. "Mind the Gap: Accounting for Measurement Error and Misclassification in Variables Generated via Data Mining," Information Systems Research (29:1), pp. 4–24.