Research in International Business and Finance 64 (2023) 101850

Contents lists available at ScienceDirect

Research in International Business and Finance

journal homepage: www.elsevier.com/locate/ribaf

Effect of twitter investor engagement on cryptocurrencies during the COVID-19 pandemic

Ahmed Bouteska [a], Petr Hajek [b], Mohammad Zoynul Abedin [c,*], Yizhe Dong [d]

[a] Faculty of Economics and Management of Tunis, University of Tunis El Manar, Tunisia
[b] Science and Research Centre, Faculty of Economics and Administration, University of Pardubice, Studentska 84, 532 10 Pardubice, Czech Republic
[c] Department of Finance, Performance & Marketing, Teesside University International Business School, Teesside University, Middlesbrough, TS1 3BX Tees Valley, United Kingdom
[d] University of Edinburgh Business School, University of Edinburgh, Edinburgh, United Kingdom

ARTICLE INFO

JEL classification:
G19
G40
G41

Keywords:
Twitter
Investor engagement
Dogecoin
Ethereum
ARIMAX
COVID-19

ABSTRACT

This study aims to examine whether the prices and returns of two cryptocurrencies, Dogecoin and Ethereum, are affected by Twitter engagement following the COVID-19 pandemic. We use the autoregressive integrated moving average with explanatory variables model to integrate the effects of investor attention and engagement on Dogecoin and Ethereum returns using data from December 31, 2020, to May 12, 2021. The results provide evidence supporting the hypothesis of a strong effect of Twitter investor engagement on Dogecoin returns; however, no potential impact is identified for Ethereum. These findings add to the growing evidence regarding the effect of social media on the cryptocurrency market and have useful implications for investors and corporate investment managers concerning investment decisions and trading strategies.

1. Introduction

Since the beginning of the pandemic, social media has been important in the cryptocurrency market because it was the only means to stay connected around the world. In the past years, social networks were largely used simply to interact with people, express opinions, and make new connections. Recently, numerous studies have been conducted addressing the effect of social media on the cryptocurrency market (Abraham et al., 2018; Li et al., 2019). Social media platforms are considered a major source of information for the cryptocurrency investors because no strong underlying fundamentals are available. Consequently, nonfundamental determinants, such as user sentiment, social media sentiment, and engagement, are expected to insignificantly impact the cryptocurrency valuation (Naeem et al., 2020, 2021; Gaies et al., 2021; Aharon et al., 2022).

There are multiple studies that empirically suggest different effects that Twitter had on cryptocurrencies before the COVID-19 pandemic (Lamon et al., 2017; Abraham et al., 2018; Dipple et al., 2020); however, a few studies have been found that provide evidence during the COVID-19 pandemic (Umar et al., 2021). Moreover, recent studies tend to emphasize the effects of tweet sentiment on the cryptocurrency market during the COVID-19 pandemic (Pano and Kashef, 2020; Corbet et al., 2020), rather than considering the investor engagement. Contrary to traditional media channels, engagement with other investors is typical for social media as it conveys certain perception of the credibility of a post (Blankespoor, 2018; Teti et al., 2019). Greater Twitter investor engagement not only

* Corresponding author.
E-mail address: m.abedin@tees.ac.uk (M.Z. Abedin).

https://doi.org/10.1016/j.ribaf.2022.101850
Received 25 April 2022; Received in revised form 12 December 2022; Accepted 17 December 2022
Available online 20 December 2022
0275-5319/

indicates mutual investor recognition and confirmation of the statement but also increases investment response to negative comments (Cade, 2018). Furthermore, Öztürk and Bilgiç (2021) highlighted that the most impactful Twitter accounts are driving the change in Bitcoin return. Hence, our study fills this immediate knowledge gap by analyzing the impact of Twitter investor engagement indicators by employing the autoregressive integrated moving average with explanatory variables (ARIMAX) models and combining them with investor attention to study their effects during the COVID-19 pandemic. The ARIMAX model was employed to incorporate the effect of explanatory variables in the cryptocurrency time series forecasting. Moreover, the ARIMAX model was selected because it was superior to deep learning-based neural networks in previous related studies in terms of prediction accuracy (Serafini et al., 2020).

This study aims to better understand and analyze how Twitter engagement conditions the cryptocurrency market, especially with respect to the second largest cryptocurrency, Ethereum and the new cryptocurrency, Dogecoin. Therefore, we download more than one year of tweets and prices for Dogecoin and Ethereum from the Twitter application programming interfaces (API) and Coinbase API via Python to study daily trading prices and returns. In particular, we implement multiple ARIMAX models to analyze the impact of exogenous Twitter investor attention and engagement variables, such as tweet volumes, retweets, replies, and likes, on the log returns of selected cryptocurrencies during the pandemic period from December 31, 2020, to May 12, 2021. The models are run subject to the requirements of stationarity condition and absence of serial correlation on the residuals. The results demonstrate a strong effect of Twitter investor engagement on Dogecoin returns, whereas it has no significant effect on Ethereum.

The results have several implications and contributions. First, this study highlights the mainstream research on the cryptocurrency trading by guiding the use of social media engagement data for predicting the cryptocurrency price. This study tends to expand the empirical insights on the relevant impacts of social media engagement that can potentially determine the context of cryptocurrency investing. Second, it demonstrates the relevance of such social media impacts on cryptocurrencies in a different era following the health crisis associated with COVID-19. Furthermore, this study can serve as a framework for the information contained in virtual communities, such as Twitter, as it had a much larger impact on the cryptocurrency markets during the COVID-19 pandemic.

The remainder of this paper is structured as follows. Section 2 reviews the relevant literature. Section 3 describes the data and outlines the methodology employed. Section 4 presents the results of the ARIMAX models. Section 5 discusses the obtained results and provides the implications. Finally, Section 6 concludes the study with directions for future research.

2. Literature review

Recently, numerous academics and researchers have examined the impact of social media on the cryptocurrency markets (Table 1). In this field, the study by Abraham et al. (2018) is considered important as it investigates the impact of social media platforms on major cryptocurrencies. The authors analyzed and predicted the price change of Bitcoin and Ethereum using data from Twitter and Google Trend. Furthermore, using a linear model, they found that Twitter volumes rather than Twitter sentiment can be used for price movement and direction. Moreover, this finding was confirmed by predicting the Bitcoin price fluctuation using neural networks with deep learning (Mittal et al., 2019). These effects were later examined for nine major cryptocurrencies by Kraaijeveld and De Smedt (2020), suggesting that Twitter sentiment is relevant for Bitcoin and Litecoin, whereas tweet volume was a significant predictor for Litecoin and XRP prices.

Colianni et al. (2015) investigated the possible use of Twitter messages and sentiment for Bitcoin to develop trading strategies. They performed various supervised learning algorithms, such as logistic regression, support vector machines, and Naïve Bayes, and found that these models improve their trading strategy by approximately 25 %. McCoy and Rahimi (2020) exploited Twitter sentiment indicators obtained using VaderSharp to analyze Japanese and English tweets, allowing them to yield a profit of 28–122 % for the testing period. Twitter sentiment and Google Trends were used by Wolk (2020) to predict the short-term cryptocurrency prices,

Table 1
Summary of previous studies on Twitter and the cryptocurrency market.

Study	Period	Twitter variables	Findings
Colianni et al. (2015)	11/2015–12/2015	volume, sentiment	Twitter data on cryptocurrencies can be used to develop accurate cryptocurrency trading strategies.
Garcia and Schweitzer (2015)	2/2011–1/2014	volume, emotional valence	Trading strategy based on Twitter signals resulted in high profits for Bitcoin.
Abraham et al. (2018)	3/2018–6/2018	volume, sentiment	Tweet volume is a more informative predictor of the Bitcoin and Ethereum price direction than tweet sentiment.
McCoy and Rahimi (2020)	11/2017–3/2018	sentiment	A lexicon-based sentiment analysis of social media combined with the relative strength index to yield high-profit trading strategies for multiple cryptocurrencies.
Dipple et al. (2020)	1/2017–8/2017	volume	Cryptocurrency market trends can be accurately forecasted using social media activity.
Choi (2021)	8/2013–5/2018	volume	Positive effect of tweet volume on Bitcoin liquidity observed for high-frequency data.
Poongodi et al. (2021)	4/2011–5/2018	topics	Topics identified to increment the prediction performance of Bitcoin price variations.
Aharon et al. (2022)	1/2013–7/2020	uncertainty	Strong causal relationship observed between uncertainty voiced in Twitter and cryptocurrency returns (especially Bitcoin).
This study	1/2021–5/2021	volume, engagement (counts of retweets, replies, and likes)	

concluding that social media platforms affect investor decisions and allow them to produce profitable prediction models.

In addition to the cryptocurrency price and return, other indicators of the cryptocurrency market, such as volatility and liquidity, were employed as response variables in previous studies. Li et al. (2019) examined Twitter signals to predict the price fluctuations of a small-cap alternative cryptocurrency called ZClassic. They used the gradient boosting tree model, and the findings demonstrated an accuracy of 80 % when used for price predictions. Furthermore, the positive effect of tweet volume (representing investor attention) on the Bitcoin liquidity was demonstrated using high-frequency data (Choi, 2021). However, the effect was only immediate, lasting only for approximately an hour.

Most recently, several studies have investigated the effect of the COVID-19 pandemic on the cryptocurrency market as the market was heavily affected. Although this market has reached new peaks, collapsed, and recovered again during the COVID-19 pandemic, the impact on individual cryptocurrencies has varied significantly (Umar et al., 2021). Dipple et al. (2020) used correlated stochastic differential equations to predict cryptocurrency prices by employing social media activity. Twitter-based uncertainty was an important indicator of Bitcoin returns during the pandemic (French, 2021; Wu et al., 2021). In particular, Bitcoin and Ethereum were reportedly net transmitters of COVID-19 shocks (Umar et al., 2021).

The aforementioned study inspired us to develop a prediction model used to analyze the effect of Twitter engagement on Dogecoin and Ethereum returns. Corresponding to previous research (Shen et al., 2019; Choi, 2021), we use the number of tweets (volume) as a proxy of the investor attention. Uniquely, we combine the investor attention with Twitter engagement by employing the numbers of retweets, replies, and likes to estimate the investor engagement in the cryptocurrency market. Furthermore, we focus on two cryptocurrencies that gained popularity during the COVID-19 pandemic, rather than relying on frequently used Bitcoin data.

3. Data and methodology

We first describe how the data for Twitter investor engagement and the two selected cryptocurrencies were collected from the Twitter and Coin Base APIs. Furthermore, we illustrate how the data were cleaned and modified for the study. Second, we explain the methodology that is employed. Cross-correlation is first performed to analyze the highest correlation periods between Twitter investor engagement and Ethereum and Dogecoin prices, subsequently the time series are checked for the stationarity condition to remove any potential bias, and finally, the impact of exogenous Twitter investor indicators on the returns of the two cryptocurrencies is analyzed and quantified using the ARIMAX models.

The first step is to collect data using a Twitter Developers account, which allows access to a full archive of historical tweets and download of 10,000,000 tweets per month at a frequency of 500 tweets per second and a maximum of 500 requests every 15 min. The account provides a bear token for the Twitter API, allowing the use of various programming languages such, as Python or RStudio, to retrieve tweets and collect them in compressed files. Moreover, the account allows the collection of data from the full Twitter search archive and tweets related to different types of queries. For instance, it is possible to retrieve tweets for a specific author or ID, select a period in which search words or hashtags were mentioned, or select language and location. For this study, specific hashtags for a period of approximately more than one year (December 31, 2019–May 12, 2021) were considered. To retrieve tweets for cryptocurrencies, the most accurate method is to search for tweets with key hashtags. Thus, #Dogecoin, #Doge, #Ethereum, and #Eth were selected. The API returned every tweet with the hashtags alongside timestamps; user IDs; and number of retweets, replies, and likes. The number of retweets, replies, and likes are other variables created from the API allowing researchers to analyze the investor engagement, indicating tweet influence as well (Zhang et al., 2018). However, searching the full Twitter archive does not provide information on the volume of tweets posted in a given period. Therefore, to analyze the number of tweets posted in a certain period, a column for these tweets was added to each collected data to obtain the volume of tweets. The final Twitter sample comprised 3,406,317 tweets with a frequency per second for the period from 31 December, 2019, to 12 May, 2021.

Subsequently, to analyze the impact of Twitter on cryptocurrency returns, the cryptocurrency historical prices were collected. For this, the Coinbase API was used to gather historical prices for Dogecoin and Ethereum. Coinbase allows free downloads of historical cryptocurrency prices as long as its API is used. The application programming interface allows collecting historical prices for Dogecoin and Ethereum covering our sample period. Sample data were downloaded on an hourly basis. Since cryptocurrencies are traded in the over-the-counter market, that is, cryptocurrency trading, which are nonpublic trades for buying or selling cryptocurrencies, they do not include opening and closing prices. To get the specific price at which we analyze the data, the adjusted closing price was calculated using the following formula for the mid-price:

$$P_t^{mid} = \frac{\left(P_t^{open} + P_t^{close}\right)}{2}, \tag{1}$$

where P_t^{open} and P_t^{close} are opening and closing cryptocurrency prices, respectively. Further, a continuously compounded return is calculated based on the mid-price to analyze the performance of Dogecoin and Ethereum, as follows:

$$R_t = \log\left(\frac{P_t^{mid}}{P_{t-1}^{mid}}\right) \tag{2}$$

To obtain the final dataset, four different datasets were combined into two datasets based on the name of the cryptocurrencies. Accordingly, the Twitter datasets were rounded to hourly timestamps, and the numbers of retweets, replies, and likes, including volumes, were aggregated. Through RStudio, the datasets were merged based on time to determine trend relationships between Twitter engagement and cryptocurrency returns. The selected period ranged from December 31, 2020, to May 12, 2021, and included

7185 h of observations.

This study chose the structured steps illustrated in Fig. 1 as the appropriate methodology to examine the effect of Twitter engagement on cryptocurrency returns.

The first step is to analyze the point at which the correlation between the Twitter engagement variables and Dogecoin and Ethereum returns is the highest. The cross-correlation is performed using the following formula:

$$r_k = \frac{\sum_{i=1}^{n-k}(x_i - \overline{x})(y_{i+k} - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{3}$$

Since the Dogecoin and Ethereum data are time series, it is important to check the stationarity of the time series data. Therefore, we considered the augmented Dickey–Fuller (ADF) unit root test. Unit roots cause unpredictable results when analyzing the time series data. The ADF test is a significance test, and thus, the null and alternative hypotheses came into play, the test statistic was calculated, and the *p*-value was reported. The *p*-value was used to determine the stationarity of the data. The ADF determined the degree of the data and how strongly or weakly the time series was defined by the trend. Stationarity implies that the statistical properties of the data do not depend on time. If we assume that the data are not stationary, then we have to change them to stationary by applying certain methods, such as differencing. The null hypothesis H_0 represents a variable that contains a unit root, whereas the alternative hypothesis H_1 represents a variable that has been generated by the stationarity process. The test rejects H_0 if the *p*-value of the ADF test is less than 0.05, indicating that the process is stationary.

We chose the ARIMAX model to include additional exogenous variables as follows:

$$y_t = \beta x_t + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} - \theta_1 z_{t-1} - \cdots - \theta_q z_{t-q} + z_t \tag{4}$$

The ARIMAX can be broken into AR, I, MA, and X, where "AR" stands for autoregressive, "I" stands for integration and stationarity, "MA" denotes moving average, and "X" means exogenous variable. This equation can be simply rewritten using the backshift operator as follows:

$$y_t = \frac{\beta}{\phi(B)} x_t + \frac{\theta(B)}{\phi(B)} z_t, \tag{5}$$

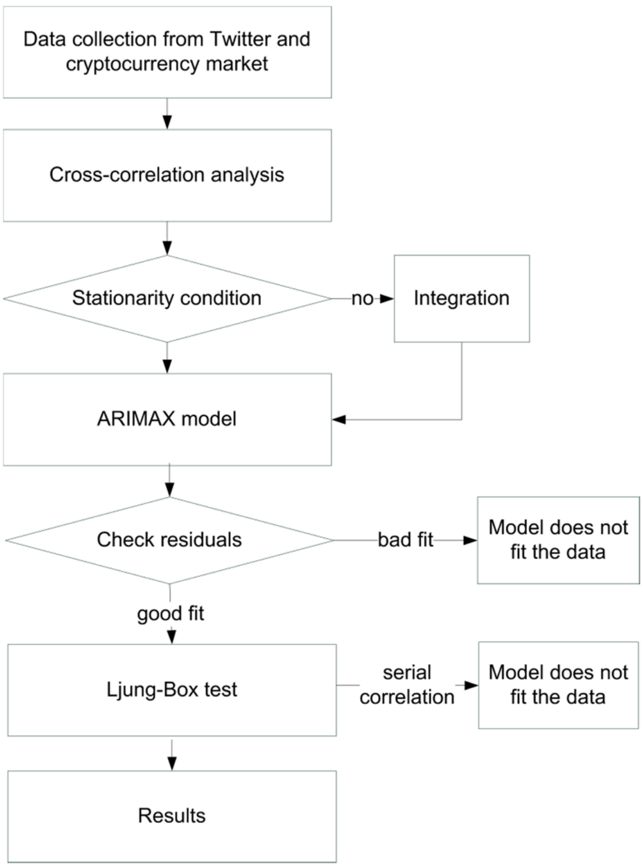


Fig. 1. Research methodology.

where, $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$.

The model includes y_t and y_{t-p}, which denote the lag time of the AR process, z_t and z_{t-q}, which denote the error terms of the MA, and x represents the exogenous variable that identifies and quantifies the effect of the external factor on the time series y. Given the time series data y_t, the advantage of the ARIMAX model is that it can be implemented to analyze past data for a given time series and allows for the prediction of future values. In defining the effect of Twitter investor attention and engagement on Dogecoin and Ethereum, we considered the following exogenous variables: Twitter volume; number of retweets, replies, and likes; and a combination of all variables together. The exogenous variable x_t can be explained as an exogenous factor affecting the endogenous variable y_t and allowing us to understand the working of the model with the new component. When looking for parameters to select to identify the best ARIMAX model, the Bayesian information criterion (BIC) was performed to select a model. The BIC can be calculated as follows:

$$BIC = -2\log likehood + \log(N) \times k, \tag{6}$$

where N refers to the number of observations and k represents the number of parameters. A small value of the BIC indicates that the model is simple with relatively fewer parameters, is the best-fitting model, and is trained on a small number of observations.

After finding the best parameters (p, d, q), we ran the ARIMAX models for various exogenous variables and examined if the models fit well by examining the autocorrelation function (ACF) and partial ACF residuals and using the Ljung–Box Q test to check for serial correlation of the residuals. We employed this test to calculate the absence of serial autocorrelation of residuals for different lag periods. Therefore, we considered the null and alternative hypotheses as follows: the null hypothesis of the Ljung–Box Q test implies that the data are independently distributed, that is, the correlations in the sample are assumed to be 0 so that any observed correlations are due to the randomness of the sampling process, that is, if the residuals are white noise. The alternative hypothesis implies that the data are not independently distributed, that is, they exhibit serial correlation. A significant p-value in this test rejects the null hypothesis that the time series are not autocorrelated. Essentially, H_0 means that the model does not exhibit poor fit, whereas H_1 means that it does.

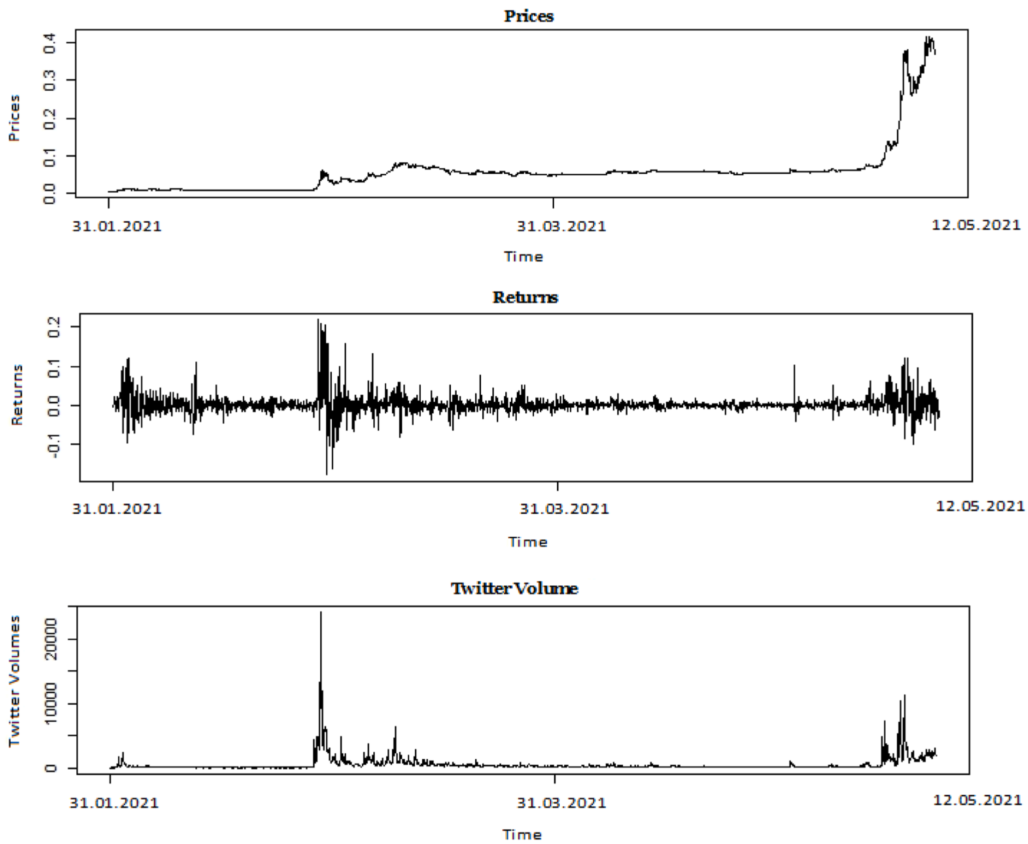


Fig. 2. Twitter volume and Dogecoin price and return series.

4. Results

4.1. Cross-correlation analysis and stationarity tests

Our results indicated that the highest correlation values for Dogecoin appeared with a time lag of one, whereas for Ethereum, the time lag was two. This implies that Twitter variables related to Dogecoin were lagged one hour earlier, and variables for Ethereum were lagged two hours earlier. Regarding the Twitter volume variable, if there was an increase in tweets posted on the platform, it would affect the return of Dogecoin one hour later and Ethereum two hours later. Figs. 2 and 3 indicate the impact of Twitter volumes on cryptocurrency prices and returns. For Dogecoin, the strong influence of Twitter on prices and returns at the end of February and April is evident. This influence is due to well-known entities, such as Elon Musk (Shahzad et al., 2022), McDonald's, and Salvadoran President Nayib Bukele, tweeting and creating related posts, especially for Dogecoin during the COVID-19 pandemic, which caused multiple problems, such as increased risks and strong short-term fluctuations in the crypto market. Elon Musk and McDonald's exchanged tweets about accepting Dogecoin for payment. Following Musk and McDonald's tweet, the price of Dogecoin rose more than 7 % to $0.14, according to Coinbase. In addition, Dogecoin was increased 11 % as Tesla and the fast-food giant accepted the cryptocurrency for store merch. After the end of the most stringent preventive and control measures of the amendment (COVID-19), we have witnessed massive rumors, false information, and negative public opinions that can cause significant turmoil in the crypto market and abnormal fluctuations, thereby increasing the volatility of the crypto market. After the pandemic, these tweets affect the users because they believe social media messages from successful people, and even though they are informed about the blockchain, they post and invest money in it, impacting cryptocurrency prices and returns. This is because Dogecoin became popular in 2021 and was created by Billy Markus, a programmer from Portland, Orlando, as a joke to mock cryptocurrencies; it has a low market that is accessible to everyone. Considering Ethereum's pandemic development, tweets regarding the currency do not seem to affect returns and prices. Certainly, it is the second largest cryptocurrency regarding the market capitalization and value, second only to Bitcoin. This means that although the volume of Ethereum-related tweets is high, the currency is unaffected because it is a well-known and trending cryptocurrency with a stable capitalization and market value.

Tables 2 and 3 report the correlation coefficients of all the variables for Dogecoin and Ethereum. For Dogecoin, we observe high correlation values of 0.297 and 0.349 between Twitter volume and Dogecoin return and price, respectively. Furthermore, we observe strong correlations between Twitter investor engagement and Dogecoin's return and price between 0.186 and 0.311, respectively. However, Ethereum shows no significant correlation coefficients for the same variables. This indicates that there is a relationship

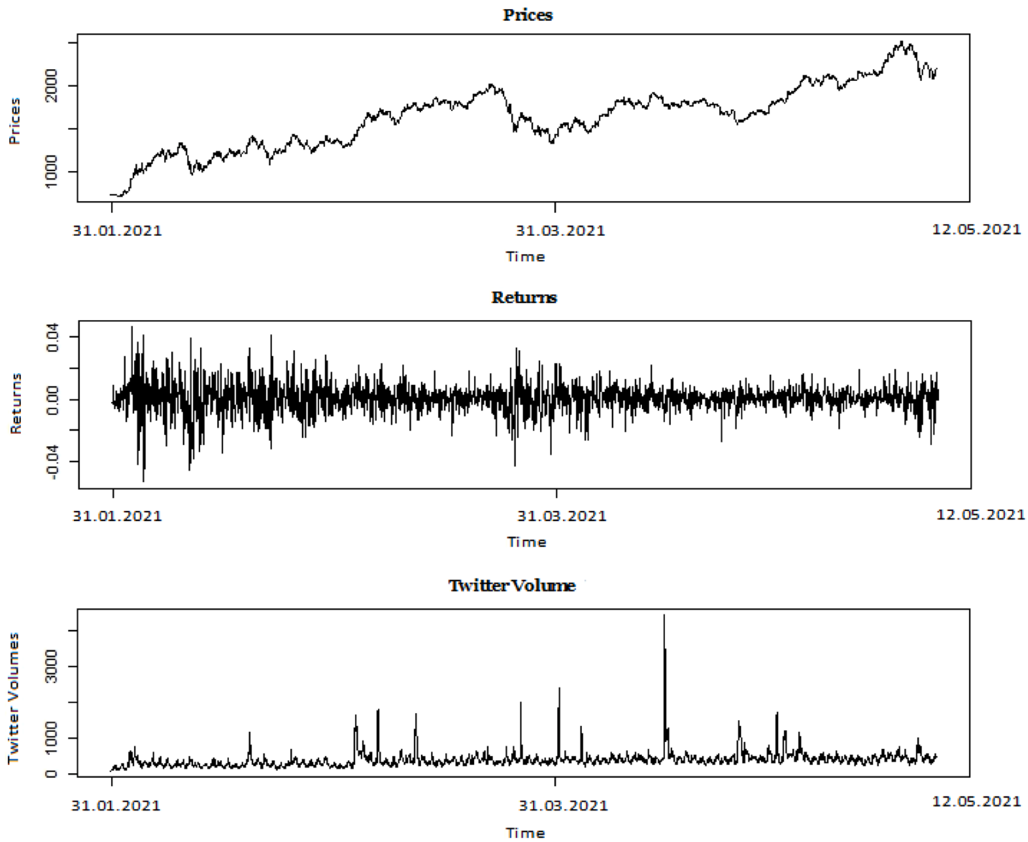


Fig. 3. Twitter volume and Ethereum price and return series.

Table 2
Correlation matrix for Dogecoin and Twitter variables.

		Return	Price	Volume	#Retweets	#Replies	#Likes	Twitter volume
Dogecoin variables	Return	1	0.044	0.296	0.226	0.186	0.228	0.297
	Price	0.044	1	0.147	0.260	0.311	0.212	0.349
	Volume	0.296	0.147	1	0.401	0.373	0.384	0.582
Twitter variables	#Retweets	0.226	0.260	0.401	1	0.823	0.908	0.741
	#Replies	0.186	0.311	0.373	0.823	1	0.957	0.700
	#Likes	0.228	0.212	0.384	0.908	0.957	1	0.713
	Twitter volume	0.297	0.349	0.582	0.741	0.700	0.713	1

Table 3
Correlation matrix for Ethereum and Twitter variables.

		Return	Price	Volume	#Retweets	#Replies	#Likes	Twitter volume
Ethereum variables	Return	1	− 0.2056	− 0.037	0.033	0.029	0.022	0.036
	Price	− 0.206	1	− 0.191	0.223	0.177	0.172	0.287
	Volume	− 0.037	− 0.191	1	− 0.024	− 0.004	0.004	0.018
Twitter variables	#Retweets	0.033	0.223	− 0.024	1	0.717	0.613	0.164
	#Replies	0.029	0.177	− 0.004	0.717	1	0.469	0.143
	#Likes	0.022	0.172	0.004	0.613	0.469	1	0.287
	Twitter volume	0.036	0.287	0.018	0.164	0.143	0.287	1

between Twitter investor engagement and Dogecoin, suggesting that an increase in the influence of tweets will only positively affect prices and returns.

Table 4 reports the results of the ADF test. We can see that all Twitter variables associated with Dogecoin and Ethereum are stationary. This means that the processes do not contain any trend and seasonality patterns and do not need to be integrated.

To investigate the causal relationship between Twitter investor engagement and the two cryptocurrencies, we adopted the methodology of Li et al. (2021a), (2021b) and used a nonparametric NWGC (wavelet-based Granger causality) test due to its robustness to heterogeneity and market complexity. Table 5 shows detailed results of the NWGC tests for Dogecoin, indicating bidirectional causalities for all the variables that represent Twitter investor engagement, including their combination. Likewise, the results in Table 6 manifest bidirectional causalities between individual variables of Twitter investor engagement and the Ethereum return, but only a unidirectional causality from the return to investor engagement. Overall, our results are consistent with those of Shen et al. (2019), Li et al. (2021a) and Tong et al. (2022), providing additional evidence on the causal relationship between Twitter investor engagement and cryptocurrency return.

To further investigate the role of Twitter investor engagement on predicting the cryptocurrency returns, we examined their connectedness using the mean-based VAR (vector autoregressive) method (Diebold and Yilmaz, 2012; Yousaf et al., 2022). More precisely, we focused on two directional spillover measures, $FROM_{i,j}$ and $TO_{j,i}$, to show the effects between variables i and j. Table 7 reports the results of the connectedness between Twitter investor engagement and the cryptocurrency returns, indicating that the gross directional spillovers to (from) others from (to) each of the return and Twitter variables were similar for the two cryptocurrencies. For both cryptocurrency returns, the gross directional spillovers from #Likes were largest, explaining 25.3 % and 13.1 % of the error variance in the Dogecoin return and Ethereum return, respectively. Regarding net directional spillovers, the largest were from #Likes to the Dogecoin return (15.6 %) and from #Retweets to the Ethereum return (6.0 %). Overall, the results reveal high conditional connectedness from Twitter investor engagement to cryptocurrency returns, which supports the use of prediction models.

4.2. Results of the ARIMAX Models

The best parameters (*p*, *d*, and *q*) of each ARIMAX model were selected by testing different models and identifying the one exhibiting the lowest BIC values (Table 8).

As seen in Figs. 4 and 5, the graphs of the ACF functions indicate that no autocorrelation of residuals for all five ARIMAX models with different exogenous variables.

Table 9 shows the *p*-values of the Ljung–Box Q test for different ARIMAX models. We can conclude that the test rejects the null

Table 4
Results of unit root tests for Twitter variables.

Dogecoin	ADF statistics	*p*-value	stationary	Ethereum	ADF statistics	*p*-value	stationary
Twitter volume	− 7.419	0.0004	yes	Twitter volume	− 8.171	0.0023	yes
#Retweets	− 3.691	0.0080	yes	#Retweet	− 4.426	0.0004	yes
#Replies	− 3.291	0.0039	yes	#Reply	− 4.027	0.0082	yes
#Likes	− 2.813	0.0097	yes	#Like	− 3.252	0.0056	yes

Table 5
Results of causalities for Dogecoin.

Variable	Causal relationship	*IE → R*		*R → IE*	
		Frequency (cycle per trading day)	Cycle length (trading days)	Frequency (cycle per trading day)	Cycle length (trading days)
Twitter volume	*IE ↔ R*	0.37	4	0.05–0.06, 0.08–0.11, 0.27↑	14–16, 4
#Retweets	*IE ↔ R*	0.02, 0.06, 0.11, 0.19↑	50, 16, 9, 7↓	0.05, 0.09–0.11, 0.21↑	20, 9–11, 6↓
#Replies	*IE ↔ R*	0.20↑	7↓	0.02↓, 0.05–0.06, 0.09–0.11, 0.23↑	0.50↑, 16–20,9–11, 5↓
#Likes	*IE ↔ R*	0.02, 0.12↑	50, 10↓	0.05, 0.10–0.12, 0.22↑	20, 10–12, 7↓
Combination	*IE ↔ R*	0.10, 0.16↑	10, 5↓	0.05, 0.09–0.11, 0.20↑	20, 9–11, 5↓

Notes: This table shows the results of the NWGC tests for causalities between the variables representing Twitter investor engagement (IE) and Dogecoin, where R is the Dogecoin return, IE ↔ R represents bidirectional causality, IE → R and R → IE denote unidirectional causalities, and ↑ (↓) represent equal and higher (equal and lower) than frequency and cycle length.

Table 6
Results of causalities for Ethereum.

Variable	Causal relationship	*IE → R*		*R → IE*	
		Frequency (cycle per trading day)	Cycle length (trading days)	Frequency (cycle per trading day)	Cycle length (trading days)
Twitter volume	*IE ↔ R*	0.07–0.08, 0.17–0.26	12–14, 4–6	0.01–0.02, 0.05–0.07	50–100, 14–20
#Retweets	*IE ↔ R*	0.06, 0.11–0.12	16, 8–9	0.05–0.06, 0.09–0.11, 0.14–0.16, 0.40	16–20, 9–11, 6–7, 4↓
#Replies	*IE ↔ R*	0.02↓,0.07–0.08, 0.17–0.28	0.50↑, 12–14, 4–6	0.05, 0.10, 0.18	20, 10, 6
#Likes	*IE ↔ R*	0.01↓, 0.04, 0.07, 0.10, 0.24–0.15	100↑, 27, 16, 10, 5–6	0.01↓, 0.05, 0.18	100↑, 20, 8
Combination	*R → IE*	–	–	0.05–0.06, 0.08–0.17, 0.23	16–20, 6–12, 3

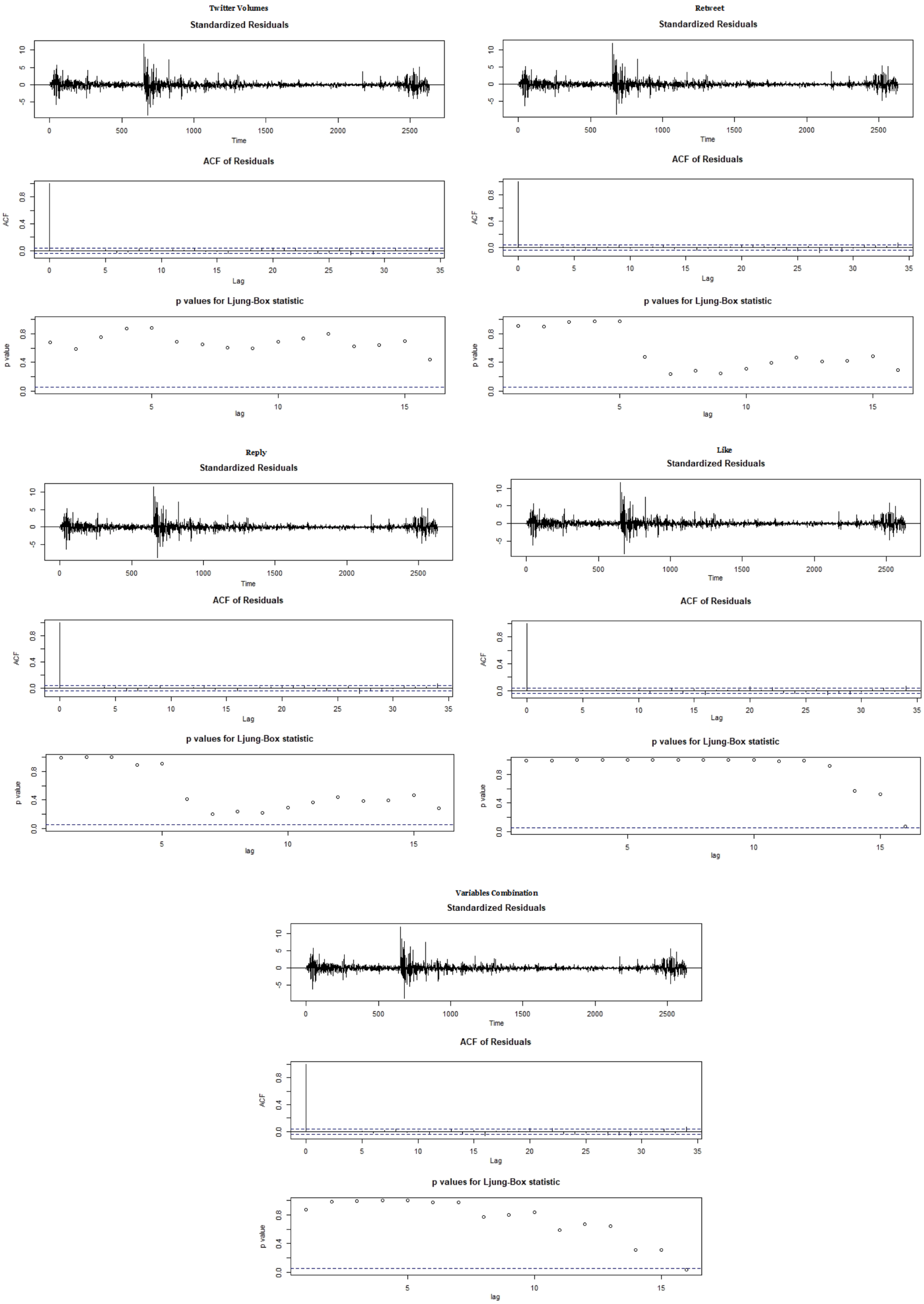
Notes: This table shows the results of the NWGC tests for causalities between the variables representing Twitter investor engagement (IE) and Ethereum, where R is the Ethereum return, IE ↔ R represents bidirectional causality, IE → R and R → IE denote unidirectional causalities, and ↑ (↓) represent equal and higher (equal and lower) than frequency and cycle length.

Table 7
Results of volatility spillovers between Dogecoin/Ethereum return and variables representing Twitter investor engagement.

	Dogecoin	Twitter volume	#Retweets	#Replies	#Likes	Combination	FROM others
Dogecoin	39.576	16.208	11.683	4.993	25.258	2.282	10.071
Twitter volume	4.587	33.180	13.214	17.802	18.583	12.633	11.137
#Retweets	7.956	23.383	21.791	11.555	26.511	8.804	13.035
#Replies	4.889	28.341	10.120	32.600	15.895	8.155	11.233
#Likes	9.685	18.745	15.346	7.886	42.274	6.063	9.621
Combination	1.413	14.931	2.696	9.452	4.732	66.775	5.537
TO others	4.755	16.935	8.843	8.615	15.163	6.323	60.634
	Ethereum	Twitter volume	#Retweets	#Replies	#Likes	Combination	FROM others
Ethereum	62.887	6.642	12.738	3.498	13.127	1.108	6.019
Twitter volume	3.950	32.438	20.919	25.369	16.246	1.079	10.927
#Retweets	6.764	22.537	34.740	13.286	21.642	1.032	10.710
#Replies	2.700	29.980	14.062	39.471	12.699	1.089	9.755
#Likes	7.575	18.579	23.254	12.302	37.240	1.050	10.293
Combination	1.007	2.737	2.365	1.164	2.547	90.18	1.470
TO others	3.166	13.412	12.223	9.270	11.043	0.060	49.174

Table 8
BIC criterion values for the ARIMAX model orders (p, d, and q).

Dogecoin	ARIMAX	BIC	Ethereum	ARIMAX	BIC
Twitter volume	(4, 0, 5)	− 14,236.37	Twitter volume	(3, 0, 5)	− 19,746.65
#Retweets	(4, 0, 5)	− 14,127.64	#Retweets	(4, 0, 5)	− 19,740.18
#Replies	(3, 0, 5)	− 14,092.43	#Replies	(3, 0, 6)	− 19,744.87
#Likes	(3, 0, 8)	− 14,092.19	#Likes	(3, 0, 3)	− 19,750.06
Combination of all variables	(3, 0, 9)	− 14,101.12	Combination of all variables	(3, 0, 3)	− 19,749.03



(caption on next page)

Fig. 4. Autocorrelation function (ACF) and partial ACF (PACF) plots of residuals for Twitter variables related to Dogecoin.

hypothesis because the *p*-values are greater than 0.05, indicating the absence of autocorrelation at the residuals for both Dogecoin and Ethereum.

After testing the correlation and stationarity conditions of the data, we met all the requirements for good model fit. Subsequently, we ran the ARIMAX model on five different Twitter variables for the cryptocurrencies (Tables 10 and 11). We noted that the root mean square error (RMSE) represents the standard deviation of the residuals, which helps determine whether the residuals are spread around the best-fitting model. In particular, the lower this value is, the better the residuals are concentrated on the ARIMAX model line. Tables 10 and 11 reveal that the RMSE values for Dogecoin and Ethereum are low, indicating good model fit. The standard errors, which represent the variance of the predicted data on the actual data, show low values, indicating that the data have good accuracy. Furthermore, we analyzed the statistical significance of the models using *p*-values, and if they are less than 0.05, it means that the models are statistically significant. The analysis aimed to measure whether the exogenous variable x affects the endogenous variable y. We found that all Twitter variables were statistically significant for Dogecoin, indicating a clear influence of Twitter investor attention and engagement on prices and returns during the study period. However, we found that no Twitter variable was statistically significant for Ethereum, indicating no effect of Twitter on prices and returns during the study period.

To validate the effectiveness of the ARIMAX models, their predictive performance was compared with two multivariate deep learning models, LSTM (long short-term memory) and BiLSTM (bidirectional long short-term memory). Following Pintelas et al. (2020), unidirectional LSTM was used, containing two LSTM layers with 30 and 15 neurons, respectively. Similarly, consistent with Pintelas et al. (2020), two BiLSTM layers with 40 and 20 neurons were used for the BiLSTM model. The Adam optimizer with 50 epochs was used to fit the deep learning models, and a 168-h rolling window was used to estimate the modes (Chu et al., 2020). From Tables 10 and 11, it can be noted that the BiLSTM and LSTM models outperformed the ARIMAX model with all variables. Despite this, the proposed ARIMAX model was competitive for both cryptocurrencies and thus the results can be considered valid.

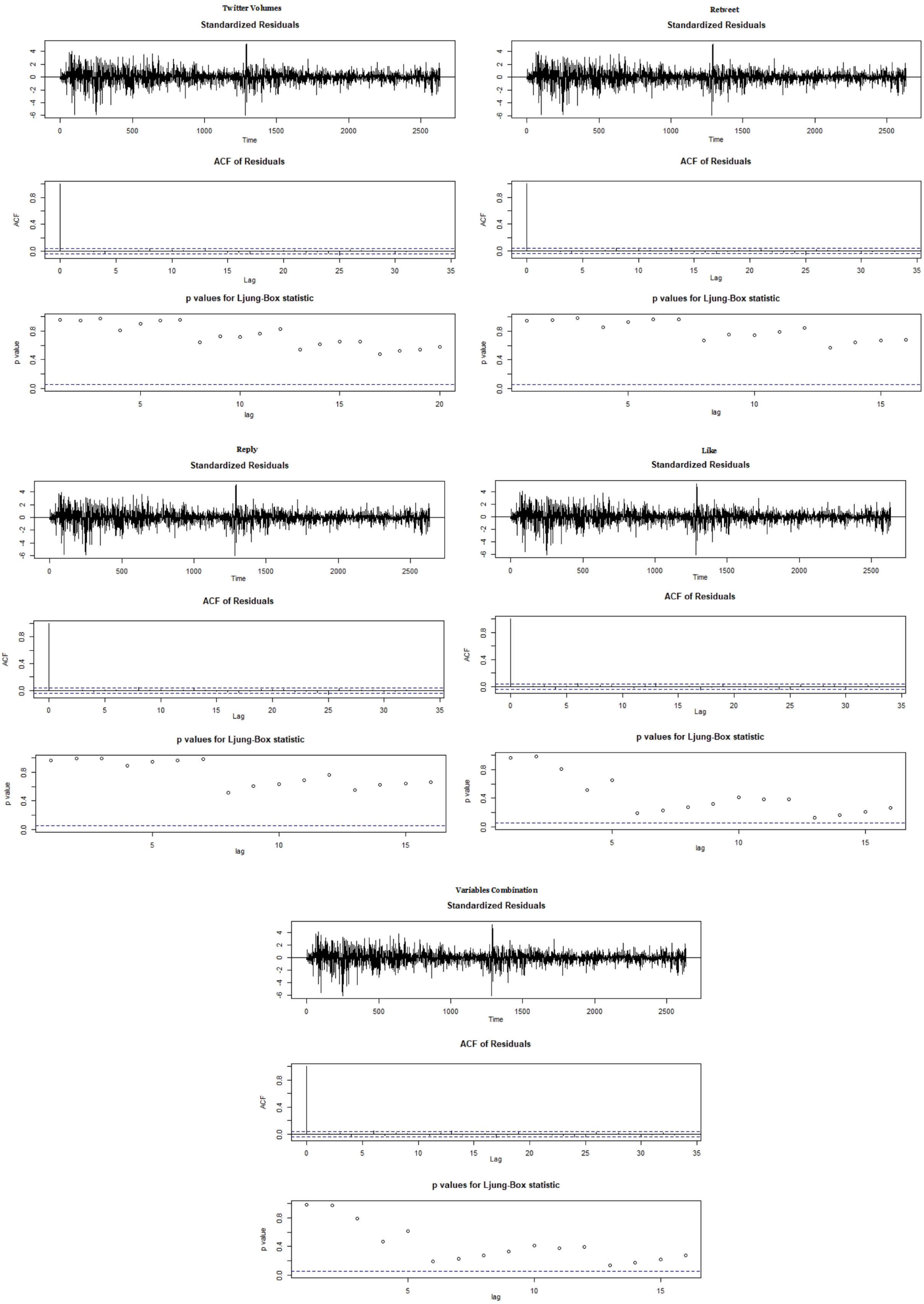
5. Discussion

The study results demonstrate strong influence of Twitter investor engagement on Dogecoin's returns but no significant effect on Ethereum. These contrasting results can be attributed to their different applications in the cryptocurrency market. The lack of investor engagement influence is likely because Ethereum, similar to Bitcoin, has a stable history and remarkable value as a safe-haven during economic crises (Mokni et al., 2022), whereas Dogecoin is a new cryptocurrency. While Ethereum is a well-known cryptocurrency with a stable capitalization, Dogecoin is highly volatile with a sharp increase in price between December 2020 and May 2021. Owing to its highly volatile prices, Dogecoin can provide high returns and has been gaining popularity, especially on social media. This has brought a second effect to the market as the more people talk about the cryptocurrencies, such as Dogecoin, the more masses start investing and prices increase rapidly. Ethereum is a decentralized application cryptocurrency in which users can interact. It is used to create non-fungible tokens, which can be traded as unique digital assets. Furthermore, Ethereum is used as a way to exchange money. Dogecoin, in contrast, originated as a joke and is mainly used for digital and peer-to-peer payments. In the late 2020s, Dogecoin became well known and gained a lot of publicity around social media, and well-known public entities, such as Elon Musk, McDonald's, and President Nayib Bukele started talking about this cryptocurrency as the future for crypto (Tandon et al., 2021). Thus, Dogecoin was noticed and well known by investors and traders globally. Having clarified this fact, our results perfectly reflect the different behavior of Dogecoin and Ethereum. For Ethereum, our findings correspond to Shen et al. (2019) and Mokni et al. (2022), suggesting a weak effect of investor attention on cryptocurrency returns. However, simultaneously, these results contradict Abraham et al. (2018), signaling a change in the investor behavior during the COVID-19 pandemic. Certainly, major cryptocurrencies surged in 2020 owing to financial speculations triggered by low interest rates (Sarkodie et al., 2022). Subsequent upward trends were partly driven by investors who recognized the upside appeal of using major cryptocurrencies to hedge against inflation owing to fixed supplies. Our results revealed that Ethereum recovered quickly and followed the efficient market hypothesis, while Dogecoin did not, which is consistent with recent empirical findings reporting different efficiencies for different cryptocurrency clusters (Sigaki et al., 2019).

Our analysis has several implications for investors. The weak response of Ethereum price to Twitter investor attention and sentiment during the COVID-19 pandemic can be valuable to investors as Ethereum can be a safe-haven in turbulent and volatile market conditions. Even though a relatively high volatility of Bitcoin and Ethereum was observed during the pandemic, this can largely be attributed to the safe-haven effect of cryptocurrencies during pandemic uncertainty (Sarkodie et al., 2022). In contrast, Dogecoin can be recommended as a suitable investment for risk-seeking investors who should concentrate on Twitter attention and engagement when making investment decisions and developing their trading strategies. Overall, our results suggest that the two cryptocurrencies can be used to effectively diversify investors' portfolio assets.

6. Conclusion

We aimed to explore the relation between cryptocurrencies and Twitter during the COVID-19 pandemic between December 31, 2020, and May 12, 2021. The main objective was to analyze the impact of Twitter investor engagement expressed through numerous retweets, replies, and likes on the price and return of Dogecoin and Ethereum using a stochastic ARIMAX model. The model results provide evidence that for Dogecoin, there is a significant effect of all the Twitter engagement variables on price and return. In contrast, the ARIMAX models for Ethereum revealed no significant effects.



(caption on next page)

Fig. 5. ACF and PACF plots of residuals for Twitter variables related to Ethereum.

Table 9
Results of Ljung–Box Q tests for Twitter variables.

Dogecoin	*p*-value	serial correlation	Ethereum	*p*-value	serial correlation
Twitter volume	0.709	no	Twitter volume	0.657	no
#Retweets	0.508	no	#Retweets	0.682	no
#Replies	0.474	no	#Replies	0.653	no
#Likes	0.449	no	#Likes	0.225	no
Combination of all variables	0.321	no	Combination of all variables	0.232	no

Table 10
Results of the ARIMAX and deep learning models with Twitter variables for Dogecoin.

Dogecoin	ARIMAX model	RMSE	Std. error	*p*-value
Twitter volume	(4, 0, 5)	0.01932	0.0006	2.3324e−17***
#Retweets	(4, 0, 5)	0.02070	0.0004	0.00813***
#Replies	(3, 0, 5)	0.01658	0.0005	0.03928**
#Likes	(3, 0, 8)	0.03031	0.0004	0.00275***
Combination of all variables	(3, 0, 9)	0.01964	0.0002	5.8081e−07***
LSTM	–	0.01698	0.0002	
BiLSTM	–	0.01513	0.0001	

Notes: ** indicates statistical significance at $p = 0.05$ and *** at $p = 0.01$.

Table 11
Results of the ARIMAX and deep learning models with Twitter variables for Ethereum.

Dogecoin	ARIMAX model	RMSE	Std. error	*p*-value
Twitter volume	(3, 0, 5)	0.00718	1.6947e−05	0.56660
#Retweets	(4, 0, 5)	0.00726	7.5932e−06	0.71598
#Replies	(3, 0, 6)	0.00570	7.6858e−06	0.46771
#Likes	(3, 0, 3)	0.00295	8.4329e−06	0.67812
Combination of all variables	(3, 0, 3)	0.00739	3.6883e−06	0.96897
LSTM	–	0.00691	1.5730e−06	
BiLSTM	–	0.00650	0.9770e−06	

Perhaps the main limitation of the current study is that tweet sentiment or emotional valence was not considered in the proposed model. Even though previous research suggested that tweet volume is more informative than tweet sentiment for predicting cryptocurrency prices and returns (Abraham et al., 2018), several other studies indicate the role of social media sentiment for certain cryptocurrencies (Steinert and Herff, 2018). Therefore, to expand our research we intend to extend the model by performing the sentiment and emotion analysis of the collected tweets (Sailunaz and Alhajj, 2019; Hajek and Novotny, 2022). As recommendations for future research, various cryptocurrencies should be examined. Further extension of the period of observation would also be interesting, especially regarding the post-COVID-19 period. Moreover, a further study could explore the effects of Twitter investor engagement by using different linear and nonlinear models, such as deep learning-based neural networks and support vector regression (Yasir et al., 2021), to provide more definitive evidence. Finally, the effect of cryptocurrency manipulation through false rumors and presence of Twitter bots should be addressed in future research (Mirtaheri et al., 2020).

Ethics approval statement

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest disclosure

The authors declared no potential conflicts of interest.

Data Availability

Data will be made available on request.

Acknowledgement

This work has been supported by the scientific research project of the Czech Sciences Foundation Grant No. 22-22586S.

References

Abraham, J., Higdon, D., Nelson, J., Ibarra, J., 2018. Cryptocurrency price prediction using tweet volumes and sentiment analysis. SMU Data Sci. Rev. 1 (3), 1.
Aharon, D.Y., Demir, E., Lau, C.K.M., Zaremba, A., 2022. Twitter-based uncertainty and cryptocurrency returns. Res. Int. Bus. Financ. 59, 101546.
Blankespoor, E., 2018. Firm communication and investor response: a framework and discussion integrating social media. Account. Organ. Soc. 68, 80–87.
Cade, N.L., 2018. Corporate social media: how two-way disclosure channels influence investors. Account. Organ. Soc. 68, 63–79.
Choi, H., 2021. Investor attention and bitcoin liquidity: evidence from bitcoin tweets. Financ. Res. Lett. 39, 101555.
Chu, J., Chan, S., Zhang, Y., 2020. High frequency momentum trading with cryptocurrencies. Res. Int. Bus. Financ. 52, 101176.
Colianni, S., Rosales, S., Signorotti, M., 2015. Algorithmic trading of cryptocurrency based on Twitter sentiment analysis. CS229 Proj. 1 (5).
Corbet, S., Hou, Y.G., Hu, Y., Larkin, C., Oxley, L., 2020. Any port in a storm: cryptocurrency safe-havens during the COVID-19 pandemic. Econ. Lett. 194, 109377.
Diebold, F.X., Yilmaz, K., 2012. Better to give than to receive: predictive directional measurement of volatility spillovers. Int. J. Forecast. 28 (1), 57–66.
Dipple, S., Choudhary, A., Flamino, J., Szymanski, B.K., Korniss, G., 2020. Using correlated stochastic differential equations to forecast cryptocurrency rates and social media activities. Appl. Netw. Sci. 5 (1), 1–30.
French, J.J., 2021. # Bitcoin, # COVID-19: Twitter-based uncertainty and Bitcoin before and during the pandemic. Int. J. Financ. Stud. 9 (2), 28.
Gaies, B., Nakhli, M.S., Sahut, J.M., Guesmi, K., 2021. Is Bitcoin rooted in confidence?–Unraveling the determinants of globalized digital currencies. Technol. Forecast. Soc. Change 172, 121038.
Garcia, D., Schweitzer, F., 2015. Social signals and algorithmic trading of Bitcoin. R. Soc. Open Sci. 2 (9), 150288.
Hajek, P., Novotny, J., 2022. Fuzzy rule-based prediction of gold prices using news affect. Expert Syst. Appl. 193, 116487.
Kraaijeveld, O., Smedt, J, De, 2020. The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. J. Int. Financ. Mark. Inst. Money 65, 101188.
Lamon, C., Nielsen, E., Redondo, E., 2017. Cryptocurrency price prediction using news and social media sentiment. SMU Data Sci. Rev. 1 (3), 1–22.
Li, T.R., Chamrajnagar, A.S., Fong, X.R., Rizik, N.R., Fu, F., 2019. Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model. Front. Phys. 7, 98.
Li, Y., Goodell, J.W., Shen, D., 2021a. Comparing search-engine and social-media attentions in finance research: Evidence from cryptocurrencies. Int. Rev. Econ. Financ. 75, 723–746.
Li, Y., Goodell, J.W., Shen, D., 2021b. Does happiness forecast implied volatility? Evidence from nonparametric wave-based Granger causality testing. Q. Rev. Econ. Financ. 81, 113–122.
McCoy, M., Rahimi, S., 2020. Prediction of highly volatile cryptocurrency prices using social media. Int. J. Comput. Intell. Appl. 19 (04), 2050025.
Mirtaheri, M., Abu-El-Haija, S., Morstatter, F., Ver Steeg, G., Galstyan, A., 2021. Identifying and analyzing cryptocurrency manipulations in social media. IEEE Trans. Comput. Soc. Syst. 8 (3), 607–617.
Mittal, A., Dhiman, V., Singh, A., Prakash, C., 2019. Short-term bitcoin price fluctuation prediction using social media and web search data. In: 2019 Twelfth International Conference on Contemporary Computing (IC3). IEEE, pp. 1–6.
Mokni, K., Bouteska, A., Nakhli, M.S., 2022. Investor sentiment and Bitcoin relationship: A quantile-based analysis. North Am. J. Econ. Financ. 60, 101657.
Naeem, M.A., Farid, S., Faruk, B., Shahzad, S.J.H., 2020. Can happiness predict future volatility in stock markets? Res. Int. Bus. Financ. 54, 101298.
Naeem, M.A., Mbarki, I., Suleman, M.T., Vo, X.V., Shahzad, S.J.H., 2021. Does Twitter happiness sentiment predict cryptocurrency? Int. Rev. Financ. 21 (4), 1529–1538.
Öztürk, S.S., Bilgiç, M.E., 2021. Twitter & bitcoin: are the most influential accounts really influential? Appl. Econ. Lett. 1–4. https://doi.org/10.1080/13504851.2021.1904104.
Pano, T., Kashef, R., 2020. A complete VADER-based sentiment analysis of bitcoin (BTC) tweets during the era of COVID-19. Big Data Cogn. Comput. 4 (4), 33.
Pintelas, E., Livieris, I.E., Stavroyiannis, S., Kotsilieris, T., Pintelas, P., 2020. Investigating the problem of cryptocurrency price prediction: a deep learning approach. In: IFIP International Conference on Artificial Intelligence Applications and Innovations. Springer, Cham, pp. 99–110.
Poongodi, M., Nguyen, T.N., Hamdi, M., Cengiz, K., 2021. Global cryptocurrency trend prediction using social media. Inf. Process. Manag. 58 (6), 102708.
Sailunaz, K., Alhajj, R., 2019. Emotion and sentiment analysis from Twitter text. J. Comput. Sci. 36, 101003.
Sarkodie, S.A., Ahmed, M.Y., Owusu, P.A., 2022. COVID-19 pandemic improves market signals of cryptocurrencies–evidence from Bitcoin, Bitcoin Cash, Ethereum, and Litecoin. Financ. Res. Lett. 44, 102049.
Serafini, G., Yi, P., Zhang, Q., Brambilla, M., Wang, J., Hu, Y., Li, B., 2020. Sentiment-driven price prediction of the bitcoin based on statistical and deep learning approaches. In: 2020 International Joint Conference on Neural Networks (IJCNN). IEEE, pp. 1–8.
Shahzad, S.J.H., Anas, M., Bouri, E., 2022. Price explosiveness in cryptocurrencies and Elon Musk's tweets. Financ. Res. Lett., 102695
Shen, D., Urquhart, A., Wang, P., 2019. Does twitter predict Bitcoin? Econ. Lett. 174, 118–122.
Sigaki, H.Y., Perc, M., Ribeiro, H.V., 2019. Clustering patterns in efficiency and the coming-of-age of the cryptocurrency market. Sci. Rep. 9 (1), 1–9.
Steinert, L., Herff, C., 2018. Predicting altcoin returns using social media. PLoS One 13 (12), e0208119.
Tandon, C., Revankar, S., Parihar, S.S., 2021. How can we predict the impact of the social media messages on the value of cryptocurrency? Insights from big data analytics. Int. J. Inf. Manag. Data Insights 1 (2), 100035.
Teti, E., Dallocchio, M., Aniasi, A., 2019. The relationship between twitter and stock prices. Evidence from the US technology industry. Technol. Forecast. Soc. Change 149, 119747.
Tong, Z., Goodell, J.W., Shen, D., 2022. Assessing causal relationships between cryptocurrencies and investor attention: new results from transfer entropy methodology. Financ. Res. Lett. 50, 103351.
Umar, Z., Jareño, F., de la O González, M., 2021. The impact of COVID-19-related media coverage on the return and volatility connectedness of cryptocurrencies and fiat currencies. Technol. Forecast. Soc. Change 172, 121025.
Wołk, K., 2020. Advanced social media sentiment analysis for short-term cryptocurrency price prediction. Expert Syst. 37 (2), e12493.
Wu, W., Tiwari, A.K., Gozgor, G., Leping, H., 2021. Does economic policy uncertainty affect cryptocurrency markets? Evidence from Twitter-based uncertainty measures. Res. Int. Bus. Financ. 58, 101478.
Yasir, M., Attique, M., Latif, K., Chaudhary, G.M., Afzal, S., Ahmed, K., Shahzad, F., 2022. Deep-learning-assisted business intelligence model for cryptocurrency forecasting using social media sentiment. J. Enterp. Inf. Manag. https://doi.org/10.1108/JEIM-02-2020-0077.
Yousaf, I., Youssef, M., Goodell, J.W., 2022. Quantile connectedness between sentiment and financial markets: evidence from the S&P 500 twitter sentiment index. Int. Rev. Financ. Anal. 83, 102322.
Zhang, X., Shi, J., Wang, D., Fang, B., 2018. Exploiting investors social network for stock prediction in China's market. J. Comput. Sci. 28, 294–303.