Market Manipulation of Cryptocurrencies: Evidence from Social Media and Transaction Data

WEN LI and LINGFENG BAO, The State Key Laboratory of Blockchain and Data Security, Zhejiang University, XiHu Qu, China
JIACHI CHEN, Sun Yat-sen University, XiangZhou Qu, China
JOHN GRUNDY, Monash University, Melbourne, Australia
XIN XIA, Software Engineering Application Technology Lab Huawei, BingJiang Qu, China
XIAOHU YANG, The State Key Laboratory of Blockchain and Data Security, Zhejiang University, XiHu Qu, China

The cryptocurrency market cap has experienced a great increase in recent years. However, large price fluctuations demonstrate the need for governance structures and identify whether there are market manipulations. In this article, we conduct three analyses—social media data analysis, blockchain data analysis, and price bubble analysis—to investigate whether market manipulation exists on Bitcoin, Ethereum, and Dogecoin platforms. Social media data analysis aims to find the reasons for price fluctuations. Blockchain data analysis is used to find detailed behavior of the manipulators. Price bubble analysis is used to investigate the relation between price fluctuation and manipulators' behavior. By using the three analyses, we show that market manipulation exists on Bitcoin, Ethereum, and Dogecoin. However, market manipulation of Bitcoin is limited, and for most of Bitcoin's price fluctuations, we found other explanations. The price for Ethereum is the most sensitive to technical updates. Technical companies/teams usually hype some new concepts (e.g., ICO, DeFi), which causes a price spike. The price of Dogecoin has a high correlation with Elon Musk's X (formerly known as Twitter) activity, showing that influential individuals have the ability to manipulate its prices. In addition, the poor monetary liquidity of Dogecoin allows some users to manipulate its price.

CCS Concepts: • **Information systems** → *Information extraction*; • **Applied computing** → **Economics;**

Additional Key Words and Phrases: Blockchain, cryptocurrencies, market manipulation, empirical study

ACM Reference Format:
Wen Li, Lingfeng Bao, Jiachi Chen, John Grundy, Xin Xia, and Xiaohu Yang. 2024. Market Manipulation of Cryptocurrencies: Evidence from Social Media and Transaction Data. *ACM Trans. Internet Technol.* 24, 2, Article 8 (March 2024), 26 pages. https://doi.org/10.1145/3643812

This research/project was supported by the National Key Research and Development Program of China (2021YFB2701102), the National Science Foundation of China (62372398, 62141222, and U20A20173), and the Fundamental Research Funds for the Central Universities (226-2022-00064). J. Grundy's work was supported by ARC Laureate Fellowship FL190100035.
Authors' addresses: W. Li, L. Bao (Corresponding author), and X. Yang, The State Key Laboratory of Blockchain and Data Security, Zhejiang University, Yuquan Campus, No. 38, Zhe Da Road, Xihu District, Hangzhou, Zhejiang Province, China, 310000; e-mails: 22121297@zju.edu.cn, lingfengbao@zju.edu.cn, yangxh@zju.edu.cn; J. Chen (Corresponding author), Zhuhai Campus, Sun Yat-sen University, Tang Jia Wan, Zhuhai City, Guangdong Province, China, 519082; e-mail: chenjch86@mail.sysu.edu.cn; J. Grundy, Monash University, Wellington Rd, Clayton VIC, Australia, 3800; e-mail: John.Grundy@monash.edu; X. Xia, Software Engineering Application Technology Lab Huawei, Huawei Hangzhou Research Institute, 410 Jianghong Road, Changhe Street, Binjiang District, Hangzhou City, Zhejiang Province, China, 310052; e-mail: xin.xia@acm.org.

ACM 1533-5399/2024/03-ART8
https://doi.org/10.1145/3643812

1 INTRODUCTION

Recent years have witnessed the great success of Bitcoin [55]. It is the first and most famous cryptocurrency, created in 2008, which is secured by its underlying technology named *blockchain*. Since the famous "Bitcoin Pizza Day" when a programmer used 10,000 BTC to pay for two pizzas in May 2010 [43], Bitcoin has experienced a tremendous growth and reached its peak in November 2021, which was worth $69,000/BTC

The rapid growth of Bitcoin resulted in the boom of the cryptocurrencies market and the birth of different blockchain platforms. One of the most famous is Ethereum [11], developed by Vitalik Buterin in 2014. The advent of Ethereum addressed many limitations of Bitcoin, whose main application is transferring and storing values. Ethereum is the first platform that supports the running of smart contracts, programming artifacts deployed on the blockchain [32]. By using smart contracts, developers can easily deploy decentralized applications [69]. The great success of Ethereum made it become the world's second largest cryptocurrency. The price of Ethereum reached its peak in November 2021 ($4,800/ETH (Ether)).

However, the real value of cryptocurrencies is controversial. To prove that Bitcoin is worthless, two developers—Billy Markus and Jackson Palmer—decided to create a payment system as a "joke" [21]. They copied the source code of Bitcoin and created a famous Altcoin named *Dogecoin* [71] just by simply changing a few lines of Bitcoin code—for example, the supplement and name [1]. The irony is that Dogecoin has now gained great uptake and has become one of the top 10 biggest cryptocurrencies in terms of market capitalization today. The total market cap of Dogecoin exceeded $50 billion on April 2021. Although the total market cap of cryptocurrencies has increased dramatically in recent years, all of them still experience large price fluctuations. For example, the price of Bitcoin was about $20,000/BTC at the end of 2017 but dropped to $3,000/BTC in January 2019. It has been shown that the price of cryptocurrencies can easily be affected by influential individuals. For instance, Elon Musk, the CEO of Tesla, was a strong advocate of Dogecoin, which led to an 11,000% price soar of Dogecoin within 2 years [3]. Due to the lack of governance and the anonymity of cryptocurrency trading, it is likely that market manipulation exists in the cryptocurrencies market [45]. The market manipulation benefits the influential individuals and enables them to make unfair profits from other crypto-assets owners [46].

A previous study [17] analyzed the leaked transaction history of the Mt. Gox Bitcoin exchange and found that there was serious market manipulation in that exchange, which led to the price of Bitcoin soaring and dropping several times between 2011 and 2013. Although a study by Chen et al. [17] proved the existence of market manipulation on Bitcoin, their study still has great limitations. Specifically, their study only focuses on Bitcoin, although many users also hold other cryptocurrencies, such as ETH and Dogecoin. However, their analysis is based on the leaked exchange transaction data, which makes their method hard to perform on other cryptocurrencies. Last but not least, the data they analyzed was between 2011 and 2013, and the cryptocurrencies market has had great changes in recent years. Thus, their results seem out of date and cannot prove that the manipulation still exists today.

In this article, we focus on Bitcoin, Ethereum, and Dogecoin, as Bitcoin has the most significant market cap, Ethereum is the most popular platform to run smart contracts and has the second biggest market cap, and Dogecoin is the most popular Altcoin. To investigate whether market manipulation exists on these cryptocurrencies and how they affect their price, we perform three analyses: Social Media Data (SMD) analysis, blockchain data analysis, and price bubble analysis. *SMD analysis* aims to find the reasons for price fluctuations. In this article, we refer to everything that could be found on the Internet as SMD, including news, Facebook data, X, and blogs, among others. There is an observation that SMD (e.g., government policies and Musk's X) could affect the cryptocurrencies market, and some SMD could be manipulated by individuals or institutions.

We first highlight large price fluctuation dates within a 6-year period. Then, we search for the SMD that responds to these fluctuations on the Internet. After that, we use open card sorting [67] to classify the SMD into seven groups (e.g., policy, market). Finally, a quantitative analysis and word cloud analysis are conducted to investigate which kind of group affects the price most. During SMD analysis, we find that big companies' policy is one of the important factors affecting the price of cryptocurrencies. Some evidence collected from SMD shows the existence of inside trading and manipulation by these companies. For example, Coinbase (a famous cryptocurrencies exchange) was accused of insider trading in 2017. They bought a large amount of **Bitcoin Cash (BCH)** in advance and then announced the addition of BCH to its exchange, which led to the price soaring of BCH. Thus, to investigate whether the inside traders (manipulators) are popular with a blockchain platform, we conduct *blockchain data analysis* to discover abnormal activities and accounts by checking their trading histories. Finally, we conduct a *price bubble analysis* by using GSADF, which is usually used to evaluate the emergence and duration of bubbles based on historical data regarding currency prices [56]. The occurrence of a price bubble is closely related to market manipulation and speculation. Based on the SMD we collected, we analyze the reasons the bubble appears and prove the existence of market manipulation.

By using the preceding three analyses, we obtain a finding different from previous work [17] based on the leaked transaction history of Mt. Gox Bitcoin exchanges between 2011 and 2013. According to our results, market manipulation indeed exists on Bitcoin, Ethereum, and Dogecoin. However, the degree of market manipulation of Bitcoin is limited, and most of its price fluctuations could be explained by other factors. The price of Bitcoin is the most sensitive to governments' policies, and it is difficult to perform market manipulation for individuals or organizations on Bitcoin. Market manipulation is prevalent on Ethereum and Dogecoin, but in different ways. The price for Ethereum is the most sensitive to technical updates. Technical companies/teams usually hype some new concepts, such as ICO [37] and DeFi (decentralized finance) [60], to drive the price soar [28]. The price of Dogecoin has a high correlation with Elon Musk's X, which shows that influential individuals have the ability to manipulate its prices. In addition, Dogecoin has poor monetary liquidity—users who own a large number of Dogecoin ("whale users") also have the ability to control its price.

The key contributions of this article include the following:

— To the best of our knowledge, this is the most comprehensive study on market manipulation of cryptocurrencies and obtaining different findings compared with previous works. We conducted in-depth analyses on various topics, including the mechanisms influencing the cryptocurrency market, abnormal trading, price bubbles, and market manipulation behavior.

— We propose three different analysis methods to substantiate the presence of market manipulation in cryptocurrencies. These methodologies are applicable for analyzing other platforms as well. Our research design embraces a multidisciplinary approach, integrating perspectives from computer science, data engineering, and econometrics, which involve the exploration of SMD, scrutiny of transactions, and validation of cross models.

— We introduce innovative formula systems and classification frameworks to quantitatively assess the impact of various factors on the cryptocurrency market. By employing a systematic approach that combines quantitative and qualitative research methods, we have significantly enhanced the quality and reliability of our study.

— We conduct an analysis on the manipulation of Bitcoin, Ethereum, and Dogecoin prices. By thoroughly examining extensive transaction records and SMD, we provide the reasons leading to price fluctuations within a 6-year period, a list of abnormal accounts, and related abnormal behaviors on the three cryptocurrencies. Additionally, these data could be utilized for further analysis (e.g., price prediction). All data and analysis results (e.g., SMD, price data, and price bubble analysis) can be found on GitHub (https://github.com/MarketManipulation/Market-manipulation).

The rest of the article is organized as follows. We present a background in Section 2. In Sections 3 through 5, we present the motivation, approach, results and findings of SMD, blockchain data, and price bubble analysis, respectively. Next, we highlight the threats to validity in Section 6. Then, we introduce related works in Section 7. In Section 8, we conclude the article and present our future work.

2 BACKGROUND

In this section, we briefly introduce background information on blockchain, the cryptocurrency market, and methods applied in our work.

2.1 Blockchain and Cryptocurrency Mining

Blockchains are typically managed by peer-to-peer networks as a public distributed ledger, where nodes compete to verify transactions and validate new blocks by solving extremely complex mathematics, a process known as "mining." When a block is successfully created, the miner receives a pre-determined amount of Bitcoin or other cryptocurrencies [55]. The mining mechanism serves two purposes: first, it sustains the Bitcoin ecosystem spontaneously, and second, it is the only way available for new cryptocurrency to enter circulation. However, the math problems' difficulties are related to the characteristics of the blockchain system, such as network hash rate. Technical factors affect the efficiency and cost of mining, and the fees for mining affect the motivation of miners, both of which directly determine the supply of cryptocurrencies and directly affect the prices.

Blockchain fork is an unavoidable issue when it comes to blockchain technology [64]. According to the preceding, different parties are required to maintain the blockchain history and add the desired changes. When parties are not in agreement for reasons such as addressing security risks and delay in confirmation of transactions [58], the blockchain splits and alternative chains may emerge. Forks can result in the creation of entirely new cryptocurrencies, as well as have a significant effect on the price of existing ones. Many well-known cryptocurrencies, such as BCH and Bitcoin Gold, evolved from the original Bitcoin blockchain via a hard fork. Recently, some work has studied how the Bitcoin fork affects volatility of the Bitcoin price. It has been concluded that within 2 months after the fork occurs, the price of Bitcoin will apparently be affected [6].

2.2 Cryptocurrencies

Bitcoin. Bitcoin was the first application of blockchain technology as well as the first P2P currency. Its maturity and pioneering significance have promoted it to stand as the "base money" of the crypto market. Bitcoin acts as an intermediary between other cryptocurrency transactions, and it links cryptocurrencies to the commercial reality of the U.S. dollar, the stock market, the options market, and so on. As of today, more and more platforms and merchants are accepting Bitcoin, such as PayPal, and Tesla has announced that they will accept Bitcoin as a form of payment. The feverish popularity of Bitcoin comes with a double-edged sword. The price of Bitcoin is inherently volatile, prone to experiencing substantial fluctuations driven by speculative trading activity, geopolitical events, and others. Sapuric and Kokkinaki [62] conclusively demonstrated, through extensive collection and analysis of financial data, that the exchange rate of Bitcoin displays significantly greater volatility when compared to traditional currencies. This characteristic price instability introduces risks and uncertainties for both avid investors and ordinary users of the cryptocurrency. Kristoufek [41] utilized instrumental variables analysis to demonstrate that heightened trading volume, increased on-chain transfer value, and the intrinsic price dynamics of Bitcoin collectively contribute to a significant amplification of Bitcoin volatility. Based on the empirical findings, it is unlikely that Bitcoin's volatility will decrease significantly over the long term. The study conducted by Pichl and Kaizoji [50] examined the price of Bitcoin in relation to standard currencies and analyzed its volatility over a 5-year period [50]. They utilized various standard currency pairs to establish both theoretical and empirical boundaries for Bitcoin arbitrage opportunities.

The intrinsic value of Bitcoin and its price formation mechanism have been thoroughly investigated in multiple studies. Consistent findings from these investigations offer compelling evidence that Bitcoin indeed holds an inherent value, even in the face of a prolonged price bubble. For example, Hayes [36] estimated Bitcoin's value using a backtested production cost model, demonstrating that although it experiences volatility and bubbles, its price cannot collapse to zero due to inherent worth. Kristoufek [40] applied economic principles like the law of one price and quantity theory of money, determining that Bitcoin possesses a fundamental value. In December 2018, its price closely aligned with an estimated intrinsic value of around $3,500. Kubal and Kristoufek [42] examined the interaction between Bitcoin price and network hash rate. By studying aspects like electricity demand and environmental impact, they revealed a deterministic link and explored intrinsic value from these angles. A study conducted by Li and Wang [48] presented a theory-driven empirical analysis of the Bitcoin exchange rate, considering both technical and economic factors. The findings indicated that in the long term, economic fundamentals had a greater impact on the Bitcoin exchange rate, whereas technical factors became less influential following the closure of Mt. Gox.

Additionally, digital wallets remain vulnerable to hacking and theft. Prudent security measures and robust backup protocols are important considerations given the non-reversible nature of Bitcoin transactions and potential for sizeable financial losses associated with compromised wallets. While the resilience of decentralized blockchain networks is notable, judicious risk management remains essential for all participants navigating both the cryptocurrency's ongoing volatility as well as technical and operational issues.

Ethereum. The concept of Ethereum was conceived by Vitalik Buterin in 2013. Despite being essentially the same in their decentralized nature, Bitcoin and Ethereum are not directly competitive. This is because Ethereum has positioned itself as a programmable blockchain that can support dApps (decentralized apps) with smart contracts, whereas the Bitcoin blockchain was created exclusively to support trading functions. The varying structure and uses of these two blockchains may have led to Ethereum being more sensitive to technology upgrades and version iteration, whereas the Bitcoin blockchain focuses on guaranteeing transaction security. With the popularity of DeFi and NFT (non-fungible tokens), Ethereum has gained momentum in recent years [73].

ETH is the "gas" of the Ethereum network. Programs and services linked to the Ethereum network require computing power. Ethereum miners receive the transaction fees paid by their clients only in ETH and execute their requested transactions. Thus, Ethereum has real value in comparison to Bitcoin because its value is embedded in the value of Ethereum's ecosystem. For as long as Ethereum is in demand, ETH will be required.

In addition, another significant application of ETH at present is ICO (initial coin offering). The difference between ICO and IPO (initial public offering) is that an IPO issues and trades stocks, whereas an ICO issues and trades digital tokens. ETH's ICO projects offer an attractive alternative to other cryptocurrencies due to the direct link between Ethereum and dApps. The popularity of a number of ICOs based on the Ethereum network has largely increased the value of

Ethereum [9]. The transaction function is not designed as deeply as that of Bitcoin. Given that Ethereum is intended to be generalized, its application should maximize the concept of low-level compatibility. This eliminates the necessity for embedded advanced use cases, such as the "time lock" in Bitcoin. As well, Bitcoin utilizes a model based on unspent transaction outputs (UTXO) as opposed to the account-based model employed by Ethereum [11], which is less secure but more convenient.

Ethereum 2.0 introduces a new consensus mechanism called **Proof of Stake (PoS)**. PoS replaces the traditional **Proof of Work (PoW)** algorithm that prior Ethereum networks utilized. In PoW, miners utilize computational power to solve complex cryptographic puzzles to validate blocks and earn rewards. In contrast, PoS selects block validators based on the amount of ETH they hold and stake. This means that node selection in PoS is determined by the quantity of currency staked rather than computing power.

This PoS mechanism aims to improve network efficiency and security, as well as reduce energy consumption, relative to PoW. The core concept of Ethereum PoS is the use of validators. Validators are node operators that hold and stake a pre-determined amount of Ether. They are responsible for validating and assembling new transaction blocks, as well as participating in the blockchain consensus process. The Ether staked by validators serves as collateral, incentivizing honest behavior. If a validator acts maliciously or negligently, they risk forfeiting their stake funds. This mechanism helps ensure the integrity and security of the Ethereum network in PoS.

Altcoins. Altcoins are digital currencies that are an alternative to Bitcoin. Both Bitcoin and Altcoins are based on the same principles. Therefore, they share code and operate as peer-to-peer systems. Altcoins are digital currencies derived from Bitcoin, which share basic principles and blockchain structure as Bitcoin but have distinct characteristics of their own. Pirgaip et al. [59] define Altcoins as cryptocurrencies which emerged after Bitcoin, including Ethereum, Dash, Dogecoin, Litecoin, Peercoin, and Ripple. Similarly, Ciaian et al. [22] conducted a research on Bitcoin market price analysis, where they investigated and analyzed not only Bitcoin but also major fiat currencies, Altcoins, commodities, and securities. There are several types of Altcoins, including Stablecoins such as Tether, Memecoins based on social media jokes and puns, and utility tokens such as Ethereum, among others.

More than 14,000 cryptocurrencies were available by 2021. According to CoinMarketCap [61], Bitcoin alone accounted for more than 40% of the total crypto market cap at the time of writing this article, whereas Ethereum makes up more than 20%. The remaining market share is occupied by all other Altcoins. Taking Dogecoin as an example, it was created to mock the "cryptocoin has value," and its code is exactly the same as Bitcoin but only changing the name 'Bitcoin' to 'Dogecoin.'Incredibly, with Elon Musk constantly hyping it on X, Dogecoin's market value is directly "flying to the moon." With the popularization of Dogecoin and the influence of Elon Musk, the next Dogecoin being hyped is SHIBA COIN, which we refer to as "SHIB" and has risen hundreds of times in just a month.

2.3 Digital Exchanges

DEXs (distributed exchanges), as conceptualized in the emerging blockchain ecosystem, refer to decentralized exchanges that operate without dependence on centralized trading platforms. DEXs facilitate transactions on blockchain networks through the implementation of smart contracts.

Prominent examples of DEXs include Uniswap, Sushiswap, and Balancer—platforms that leverage smart contract technology to realize automated market making mechanisms for exchanging digital assets. In contrast to conventional centralized exchanges, DEXs are posited to afford users enhanced mobility and transparency due to their democratized and openly verifiable designs based on public blockchain ledgers. This novel construct of decentralized trading venues redistributes powers from centralized intermediaries to individual users, offering researchers a case study on the technical and economic impacts of decentralization in crypto markets.

2.4 Topic Modeling and Latent Dirichlet Allocation

The topic model is a statistical method for extracting topics from a given corpus based on machine learning or **Natural Language Processing (NLP)** [7]. Topic modeling is considered one of the most popular methods for uncovering hidden semantic structures in NLP.

Latent Dirichlet Allocation (LDA) is a popular topic modeling technique that enables sets of observations to be explained by unobserved groups that explain why some parts of the data are similar [8]. Specifically, it is theorized that each document consists of a limited number of words and each word in the document relates to its topic.

2.5 Cryptocurrency Bubble

Generally, economists describe "bubble" as deviations from normal distributions of asset prices. As the bubble expands, investors will suffer large losses, increasing financial system risks and the possibility of a financial crisis. Considering that bubbles deviate from the real economy and have no intrinsic value, they are extremely fragile and may adversely affect the development of the blockchain industry. The lack of intrinsic value for digital currency assets exacerbates this volatility. Accordingly, studies on the quantification and causes of cryptocurrency price bubbles have attracted the attention of a wide range of stakeholders, including regulators, policy makers, investors, and scholars. The existence of cryptocurrency bubbles has been demonstrated in a wide range of studies. According to Garcia et al. [31] and Kristoufel [39], previous research investigated the main Bitcoin price drivers and indicated the existence of a bubble phase.

3 SMD ANALYSIS

In this section, we investigate which kinds of SMD have the most impact on cryptocurrency price, and whether it could be controlled by individuals or institutions to manipulate the market.

3.1 Motivation

The cryptocurrency market has suffered extreme price fluctuations in recent years, which shows high correlations with some SMD [10]. For example, Dogecoin price always soared or fell when Elon Musk sent Dogecoin-related information on his X. Due to the lack of supervision, the fluctuation of cryptocurrencies might give chances for people or institutions to manipulate the market to make unfair profits [65]. Besides, the policy of governments also greatly influences the price of cryptocurrencies. For example, the Chinese government banned cryptocurrency trading, which led to the drop of Bitcoin price in May 2021. Although policy-related information could affect the price of cryptocurrencies, it is less likely that some people have enough power to utilize government policy to manipulate the cryptocurrency market [66].

3.2 Approach

Figure 1 illustrates an overview of the architecture of how we find the SMD, how we classify them into different groups, and their impact levels on cryptocurrency prices. Our method consists of three steps, e.g., data collection, SMD classification, and influence analysis. We first collect the daily price of Bitcoin, Ethereum, and Dogecoin from July 2016 to July 2022. We then highlight the dates with large price fluctuation[1] (10%). We then search related SMD through search engines and

[1]We regard 10% as a large price fluctuation, as in many stock markets (e.g., China, Singapore), when the stock price falls or rises greater than 10% in a single day, the trade will be terminated.

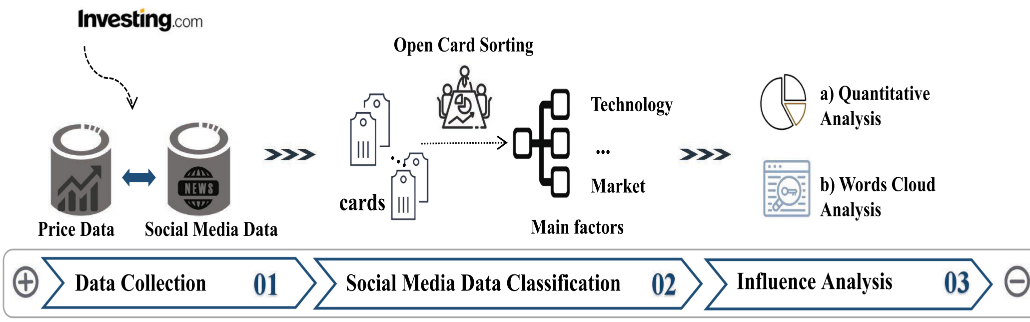

Fig. 1. SMD analysis process overview.

social media websites, such as Google, X, and Facebook. After that, we utilize open card sorting to classify the obtained SMD into seven groups. Finally, we conduct quantitative and word cloud analyses to find which factor strongly influences cryptocurrency prices. We discuss the detailed steps in the following.

It was crucial to strategically define the range of data to ensure its validity. To minimize bias, we intentionally extended the data collection period to 5 years, combining transaction data with SMD from multiple platforms. This comprehensive approach utilizing big data surpasses previous efforts that focused on narrower time windows or single sources. In comparison, Moore et al. [54] analyzed fraud data solely from the Bitcoin market during the period from 2010 to 2015. Cheung et al. [20] examined evidence of Bitcoin bubbles and manipulation using transaction logs from 2010 to 2014. In a seminal analysis published in 2023, La Morgia et al. [46] investigated pump and dump schemes across Dogecoin exchanges, taking a more comprehensive longitudinal perspective by spanning 3 years and including data from four distinct trading venues.

We took measures to increase the research scope and data integration to reduce biases, but the inherent uncertainty and rapid changes in the cryptocurrency market inevitably mean that conclusions may not remain applicable in the long run. The primary objective of this study is to address the lack of knowledge analysis in recent years and provide references for current decision making to identify risks and develop response strategies. At the same time, we established a price impact assessment framework aimed at providing a relatively stable analytical benchmark for the cryptocurrency market. The framework methodology itself is robust and provides a foundation for future research. Through continuous optimization, we believe that the factors model can better extract strategic insights from online and market systems, although the systems are full of uncertainties.

3.2.1 Data Collection. Two kinds of data need to be collected for further analysis: historical daily prices data of cryptocurrencies and SMD. The recent 6-year daily price data is collected through the API provided by Investing.com. The time period of the data we collected for the three cryptocurrencies are all between July 2016 and July 2022. After finding the fluctuation date of the three cryptocurrencies, we searched related SMD via search engines and social media websites, such as Google, X and Facebook. For example, Bitcoin price fell −15.42% on August 2, 2016. We searched related news and found that $72 million of Bitcoin was stolen from Hong Kong's Bitfinex exchange that day, which led to the drop of Bitcoin price. Table 1 shows the number of days that fall or rise greater than 10%. Specifically, there are 43 days and 32 days of Bitcoin that rise and fall greater than 10% in the 6 year period. The table shows that the Bitcoin price is the most stable, whereas Dogecoin is the worst.

Table 1. Number of Days That Fall or Rise Greater Than 10% for the Three Cryptocurrencies

Cryptocurrency	Period	R ≥ 10%	F ≤ −10%
Bitcoin	2016.07–2022.07	43	32
Ethereum	2016.07–2022.07	95	64
Dogecoin	2016.07–2022.07	140	95

Table 2. Seven Kinds of SMD Leading to Price Fluctuations in the Cryptocurrencies Market

Group	Explanation
Technology	Any protocol changes of the blockchain system (e.g., consensus protocol, performance optimization)
Policy	Governments' policy regarding the cryptocurrencies market
Market	Price fluctuation caused by specific institutions or big companies (e.g., Tesla accepts Bitcoin as a payment)
Economic	Price fluctuation caused by economic factors (e.g., financial laws and regulations, market correlations)
Security	Price fluctuation caused by attacks or any security-related issues
Views	Public opinions or actions by influencers in the cryptocurrency field (e.g., Elon Musks' X)
Others	Black swan events that lead to the fluctuation (e.g., wars, Covid-19)

3.2.2 SMD Classification. We follow the open card sorting [67] approach to analyze and categorize the collected SMD. Open card sorting is commonly used for organizing data with no predefined groups. It involves grouping fragmented elements into a grouped information architecture with inclusion relationships. Specifically, each card will be clustered into a group with a certain topic or meaning first. If there is no appropriate group, a new group will be generated.

The open card sorting approach is particularly appropriate in situations where there are many sub-information elements, and designers are not sure which sub-information elements belong to which, or even how to classify them. In addition, it organizes the elements of the information system in a way that is easy for users to understand. It can be used to help design or evaluate information architecture. Researchers may gain valuable insight from classifying cards, as well as conduct in-depth analyses to determine the principles and defining characteristics of each card.

There are two iterations of this step and two authors of this article involved in the card sorting. In *iteration 1*, we randomly chose 20% of the SMD. The two authors carefully read the details of the SMD and double-checked whether the SMD was the root cause of the price fluctuation. They then discussed to classify the SMD into a group, such as economic and policy. In *iteration 2*, the two authors independently categorized the remaining 80% of the SMD into the initial classification scheme described in iteration 1. After that, they compared their results and discussed any differences to ensure their correctness. Finally, the collected SMD were classified into seven groups; definitions of them can be found in Table 2.

3.2.3 Influence Analysis on Prices. Two kinds of analyses are conducted next: quantitative analysis and word cloud analysis. Quantitative analysis aims to find which group contributes the most to the price fluctuation. Word cloud analysis attempts to analyze the semantics of the collected SMD and find more detailed factors of the price fluctuation.

Quantitative Analysis. Based on the preceding classification, we get a list of group categories as *groups = {Technology, Policy, Market, Economic, Security, Views, Others}*. The impact of the each group is calculated as Equation (1): $group\_frequency[i]$ is represented by Equation (2), where n_i is the number of SMD classified into group i (e.g., n_1 is the number of SMD classified as Technology), and $group\_weight[i]$ is calculated as Equation (3) as follows, where m_i represents the daily rise or fall caused by an SMD, reflecting the impact of an SMD.

To be specific, following the determination of the daily fluctuations of price rises and falls, we retrieve a sizable corpus of corresponding SMD. These SMDs are prospectively organized into the pre-defined taxonomy of seven topical factors previously established through the prior card sorting process. Each SMD collected is labeled with the associated date and price fluctuation designation for said date. We utilize the price variation affiliated with each SMD to computationally derive subsequent factor weights. Subsequently, we will computationally leverage the price variation affiliated with each SMD to systematically derive the ensuing factor weights.

For example, $Technology\_News[m_1, m_2, ...]$ represents SMD in the Technology factor category, where m_i represents the price fluctuation associated with that news, indicating its impact. n_1 represents the number of SMD labeled as Technology, and n_2 represents the number labeled as Policy. The seven categories constitute the entire SMD matrix.

Next, in calculating the weights, we consider frequency of occurrence and magnitude of price fluctuation caused by a factor. We calculate the final weight factors from both $group\_frequency[i]$ and $group\_weight[i]$. Some market influences may appear frequently but not cause significant price changes, whereas other policies may significantly impact price but occur less frequently. $group\_frequency[i]$ represents occurrence frequency, and $group\_weight[i]$ represents influence level. The final weight is calculated using multiplication. If related Technology SMD do not appear on a day, their price fluctuation is considered zero, meaning that they do not contribute to the influence factor calculation.

$$group\_impact[i] = group\_frequency[i] + group\_weight[i] \tag{1}$$

$$group\_frequency[i] = \frac{n_i}{\sum_{i=1}^{7} n_i} \tag{2}$$

$$group\_weight[i] = \frac{\sum_{x=1}^{n_i} |SMD[i][m_i]|}{\sum_{x=1}^{7} \sum_{x=1}^{n_i} |SMD[i][m_i]|} \tag{3}$$

Word Cloud Analysis. We utilize the LDA topic model [7], an automated model that analyzes a combination of documents data, to create clusters of words and discover the hidden patterns between words related to their topics. Based on our experience, we summarized 5 topics for each cryptocurrency, and each topic has 30 topic words. Following the LDA analysis results, we generated a word cloud map based on the word frequency distribution for the three cryptocurrencies, which provides an intuitive view of the associated words with the three cryptocurrencies.

3.3 Results

3.3.1 Distribution of Seven SMD Groups. Table 3 shows the distribution of seven SMD groups for the three cryptocurrencies. Note that several SMD might contribute to the same day's price fluctuation. Thus, the sum of the seven factors might exceed the total days of rise/fall. For example, China imposed new restrictions on cryptocurrency trading and mining on May 19, 2021. On the same day, Tesla also announced that they would no longer accept Bitcoin as a payment. These two SMD were labeled as Policy and Market, respectively; they both contributed to the fall of the cryptocurrencies market on that day.

However, we failed to find any SMD that led to the price fluctuation on some days. For example, there are 43 days when the price of Bitcoin rose higher than 10%, but we failed to find any SMD result for the fluctuation for 9 days. The detailed number can be found in the last line of Table 3. The results show that all of the price falls of Bitcoin could be explained by related reasons. Thus, it is possible to warn about the risk to Bitcoin holders when specific events happen. The price

Table 3. Distribution of Seven SMD Groups for the Three Cryptocurrencies

	Bitcoin		Ethereum		Dogecoin	
	R(43)	F(32)	R(95)	F(64)	R(140)	F(95)
Technology	6	6	55	23	4	0
Policy	19	20	17	21	0	28
Market	19	16	45	18	55	5
Economic	10	9	22	19	25	38
Security	2	3	1	11	0	0
Views	9	3	11	8	87	19
Other	1	2	1	3	0	0
Not Found	9	0	3	4	37	29

Table 4. Importance of Seven SMD Groups for the Price Fluctuation

	Bitcoin		Ethereum		Dogecoin	
	R	F	R	F	R	F
Technology	0.004	0.010	**0.156**	0.048	0.000	0.000
Policy	**0.134**	**0.335**	0.011	**0.054**	0.000	0.078
Market	0.107	0.141	0.076	0.034	0.125	0.004
Economic	0.010	0.035	0.019	0.028	0.025	**0.185**
Security	0.000	0.003	0.000	0.011	0.000	0.000
Views	0.014	0.008	0.004	0.004	**0.222**	0.052
Others	0.000	0.001	0.000	0.001	0.000	0.000

fluctuation of Ethereum is more predictable, as most of the price fluctuation could be attributed to a specific reason. For the price fluctuations that cannot find related SMD, we found some large trading options from whale users from blockchain transactions. Selling or buying large amounts of cryptocurrencies could lead to a large price fluctuation, which might hint to the possibility of market manipulation by whale users. We discuss the details of this in the next section. Note that this kind of behavior will not be classified into the Economic group, as the time of the fluctuation does not follow any financial laws and regulations.

3.3.2 Quantitative Analysis Results. Table 4 illustrates the importance of seven SMD groups for price fluctuation by using the method we introduced in Section 3.2.3. Groups with a higher score means that they contribute more to the price fluctuation. Among the seven groups, View and Market SMD could be controlled by influential individuals or institutions, which hints at a high possibility of market manipulation. Technology and Policy are usually controlled by blockchain teams and governments, which shows a low possibility of market manipulation. Thus, we call *View* and *Market* human-controlled groups, and we call *Technology* and *Policy* non-human-controlled groups. For the other three factors, we called them *neutral factors*, as it is not easy to determine whether a specific individual/institution could control the SMD. For example, Economic is a collective action by many cryptocurrency investors. There is usually a large price fall followed by a large price rise, as many investors want to hold the profits they earned. It is likely that the price fluctuation is a consensus for all investors, but it is also possibly controlled by influential individuals, such as whale users, as they have the ability to cause panic and greed.

SMD groups contributing to price fluctuation of the three cryptocurrencies are significantly different. The *Policy* is the most important group that leads to the price fluctuation for Bitcoin. This means that the governments' attitude has the highest influence on the Bitcoin market. Ethereum is the first and the most popular blockchain platform that supports the running of smart contracts. As

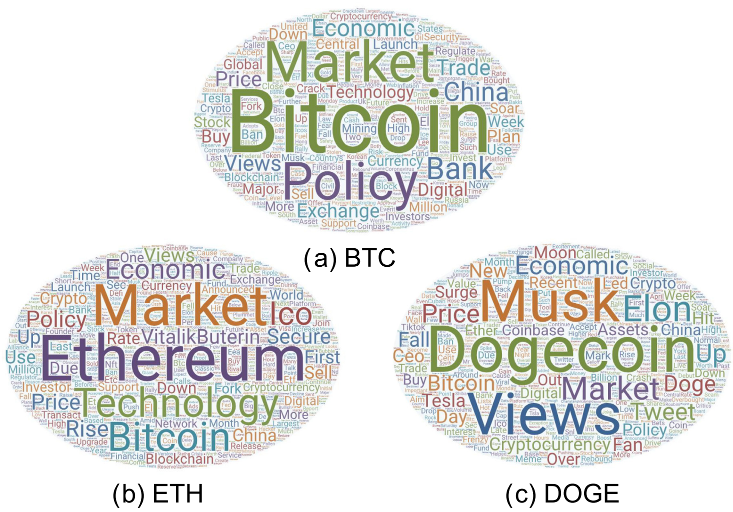

Fig. 2. Word cloud graph based on the LDA results.

expected, the *Technology* group affects ETH price the most. Although the most important groups that lead to price fluctuation for Bitcoin and Ethereum are non-human-controlled factors, one of the human-controlled groups, *Market*, also contributes a lot. This shows that big companies' attitudes also have strong influence on the price of Bitcoin and Ethereum.

Unlike Bitcoin and Ethereum, more human-controlled factors are involved in the price fluctuation of Dogecoin. Specifically, *Views* contributes much more to the price rise than other factors, whereas *Economic* contributes the most to the price fall. The non-human-controlled factors (i.e., *Technology* and *Policy*) almost have no influence on Dogecoin's price rise. This clearly shows that the price rise of Dogecoin is human controlled. Many investors clearly know this, and thus price falls are usually followed by a rise (*Economic*), as investors want to lock their profits.

3.3.3 Word Cloud Analysis Results. Following LDA analysis, we generate word cloud maps based on the frequency of topic words distribution for the three currencies. This provides an intuitive view of the associated words with each of the three currencies shown in Figure 2.

According to the word cloud map, price fluctuations of Bitcoin are strongly related to "policy," followed by "banks," "China," "market," and so on. Note that "Elon Musk" plays an important role in price fluctuations for Dogecoin, even more important than "policy" and "economics." This shows that Elon Musk has a strong influence on Dogecoin's price. "Technology" is the most important word for Ethereum, and follows with "policy," "markets," and "finance." "Vitalik Buterin," the co-founder of Ethereum, also appears in the word cloud map of Ethereum. Unlike Musk, Buterin is a developer who usually announces the technical updates/future maps of Ethereum, and rarely shows an opinion on the cryptocurrencies' price. However, it also shows that Buterin has the ability to affect the Ether price.

3.4 Findings

Based on the preceding analysis, we summarize the following findings:

— *Finding 1*: Most of the price fluctuations of the three cryptocurrencies can find related reasons from SMD. Dogecoin contains the highest number of suspicious price fluctuations, whereas

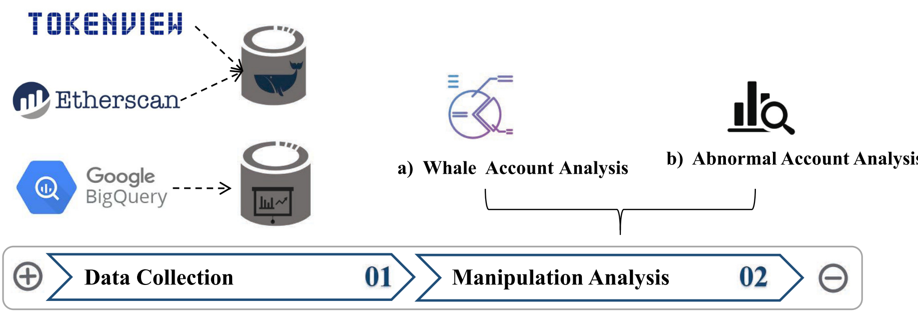

Fig. 3. Overview of the architecture of blockchain data analysis.

the Bitcoin and Ethereum price are more predictable, which hints to lower probability of market manipulation in these two cryptocurrencies.

— *Finding 2*: Bitcoin price has a high correlation to governments' policy, especially for China. The technology factor plays an important role in the Ethereum price fluctuation. The price rise of Dogecoin is seriously human controlled.
— *Finding 3*: Individuals have have little responsibility in the price fluctuation of Bitcoin and Ethereum in the our 6-year period. However, big companies and institutions' attitudes can lead to their price fluctuation.
— *Finding 4*: Influential individuals have the ability to affect the price of Ethereum and DOGE (e.g., Vitalink for Ethereum and Musk for Dogecoin). However, there is no specific person who has enough influence to affect the Bitcoin price. This indicates that manipulating the Bitcoin price is more difficult than for Ethereum and Dogecoin.

4 BLOCKCHAIN DATA ANALYSIS

In this section, we investigate whether insider traders (manipulators) are influential on each blockchain platform. Finding the blockchain accounts of such insider traders and analyzing their trading activities may allow us to analyze how they manipulate the cryptocurrency price.

4.1 Motivation

In the previous section, we found some suspicious large price fluctuations for which we cannot find any reasons. According to our observation from blockchain transactions, whale users' trading activities might take responsibility for these fluctuations. Besides, we also found that "Market" and "Views" are the two important SMD groups that result in the large price fluctuation of cryptocurrencies. However, "Market" and "Views" could be controlled by some people or institutions. We indeed found some evidence to prove that the manipulation exists. For example, Binance, the world's largest cryptocurrency exchange, was investigated for insider trading by the CFTC (the U.S. Commodity Futures Trading Commission) and the SEC (the U.S. Securities and Exchange Commission) in 2021. These SMD confirmed that whale users' behavior could affect the price of cryptocurrencies and the existence of insider trading.

4.2 Approach

Since all activities on the blockchain are stored and visible to the public, we utilize blockchain data to analyze whether manipulation and insider trading really exist. The detailed steps are illustrated in Figure 3.

Table 5. Percentage of Coins Held by Top N Whale Users

Top N	10	30	50	100
BTC	6.30%	10.37%	12.53%	15.72%
ETH	22.20%	28.51%	31.91%	38.46%
DOGE	42.90%	51.33%	55.33%	60.29%

4.2.1 Data Collection. In the data collection step, we use Tokenview[2] and Etherscan[3] to collect blockchain account information. Both of them are blockchain search engines to look up the detailed information of a specific account, such as its balance, transaction information, and identity information (although blockchain account is anonymous, some accounts will be given a name tag, e.g., "exchange" or "attacker"). Google BigQuery[4] is a serverless, cost-effective, multi-cloud data warehouse designed to provide business insights from big data and address analytics and data science problems [33]. Its data warehouse stores blockchain data of Bitcoin, Ethereum, and Dogecoin. It enables scalable analysis over petabytes of blockchain data by using SQL language.

4.2.2 Manipulation Analysis. Two analyses are conducted in this section: whale account analysis and abnormal account analysis.

Whale Account Analysis. This analysis aims to find the impact level of whale accounts on the prices of cryptocurrency. Whale accounts hold a large number of cryptocurrencies. Their trading behavior could lead to price fluctuations. We utilize Google BigQuery to find the top 100 richest whale accounts for the three cryptocurrencies and then manually check their historical trading activities.

Abnormal Account Analysis. This analysis aims to prove that inside traders and manipulation exist on the cryptocurrency market and analyze abnormal accounts' behavior. Specifically, we first conduct a transaction analysis to find "abnormal transactions" and "abnormal accounts." A transaction is identified as an *abnormal transaction* if this transaction has "a large sized transfer/an usually large size of transfer"[5] within 5 days before a large price fluctuation, and all amounts are transferred to or from exchange accounts. The sender of the an abnormal transaction is regarded as *abnormal accounts.*

4.3 Results

4.3.1 Whale Account Analysis. Table 5 gives the percentage of cryptocurrencies held by the top N whale users. All three cryptocurrencies show a degree of centralization but still have significant differences. The top 10 whale users hold 42.90% of Dogecoin, and more than 60% of Dogecoin is owned by the top 100 users. These numbers clearly show that the liquidity of Dogecoin is poor and whale users could easily control Dogecoin prices. Ethereum also shows great centralization with the top 10 whales holding 22.20% of ETHs and the top 100 users controlling more than one-third of the total ETHs. The assets of Bitcoin are the most distributed among the three cryptocurrencies. The top 10 whale users also control 6.30% of Bitcoin, which means that it is much more difficult for Bitcoin whales to manipulate the price compared to the other two cryptocurrencies.

4.3.2 Abnormal Account Analysis. Based on the approach introduced in Section 4.2.2, we find a total of 5, 27, and 23 abnormal accounts before large price falls, and 0, 66, and 0 abnormal accounts

[2] https://tokenview.com/
[3] https://etherscan.io/
[4] https://cloud.google.com/
[5] A weekly report from Huobi [75], the top five cryptocurrency exchange, defines "a large amount of transfer"as 1,000 BTC, 10,000 ETH, or 1,000,000 DOGE in a transaction.

Fig. 4. An example of an abnormal account (DRSqEwcnJX3GZWH9Twtwk8D5ewqdJzi13k) of Dogecoin.

before large price rises for Bitcoin, Ethereum, and Dogecoin, respectively. The detailed information of these "abnormal accounts" can be found in our GitHub repository.

Figure 4 gives an example of an abnormal account of Dogecoin. The account is labeled as an abnormal account as we find that it generates several "abnormal transactions." Then, we manually check all of its trading activities. From the figure, we can see a high overlap between the account's trading activity and Dogecoin's price fluctuation. Specifically, the account always buys Dogecoin (coin transfer into the account) when the price is low and sells Dogecoin (coin transfer out of the account) when the price reaches its peak. Another interesting "coincidence" is that there are always "Market" or "View" SMD followed by its trading activities, and these SMD are regarded as the reasons leading to large price fluctuations. This coincidence leads to a possibility of insider trading and manipulation of Dogecoin.

Bitcoin has the least number of "abnormal accounts," which shows that market manipulation is rarer on Bitcoin compared to Ethereum and Dogecoin. We found that only Ethereum has abnormal accounts before a large price rise, but all three platforms find related abnormal accounts before a large price fall. Normally, selling cryptocurrencies could lead to a price drop. To sell coins, the coin owners need to transfer the cryptocurrencies to exchanges, which will leave transactions on the blockchain. Similarly, buying cryptocurrencies can increase the price value. However, transferring cryptocurrencies from exchanges to blockchain will induce fees. Thus, many coin owners might only trade on centralized exchanges, which will not generate transactions on the blockchain. The number of abnormal accounts for Ethereum before large price rises might show that the ETH owners are more confident about the future of Ethereum and more willing to have long-term holding of ETH. To our surprise, Ethereum has the highest number of abnormal accounts. We observed that many abnormal accounts only have a single abnormal transaction, and these transactions were generated in a same time period. This hints to these accounts belonging to the same institution or individual.

4.4 Findings

Based on the preceding analysis, we summarize the following findings:

— *Finding 5*: Dogecoin and Ethereum show great centralization, especially for Dogecoin. The bad monetary liquidity gives the ability to whale users to manipulate the price.

Manipulating the Bitcoin price is the most difficult among the three cryptocurrencies due to the decentralized asset distribution.

— *Finding 6*: All three cryptocurrencies appear to contain "abnormal accounts" able to "predict" the future price. The abnormal accounts prove the existence of insider trading and manipulation of the cryptocurrencies market, but only a limit number of abnormal accounts are found on Bitcoin, which shows that market manipulation is not common in the Bitcoin market.

5 PRICE BUBBLE ANALYSIS

In this section, we investigate the price bubble phenomenon on the three cryptocurrency platforms to identify the possibility of market manipulations.

5.1 Motivation

Cryptocurrency prices can deviate significantly from their real values due to anonymity and lack of regulation from government. The phenomenon is often referred to as the *bubble* [72], which is closely related to market manipulation and speculation.

Based on SMD collected over the past few years, it appears that a small number of addresses are manipulating the market by leveraging their social influence and crypto-assets, profiting by repeatedly triggering price explosions and hyping cryptocurrencies. By detecting price bubbles, we can identify periods of market manipulation and locate abnormal trading activity. As outlined in the previous section, we will compare the duration of three types of cryptocurrency bubbles with the duration of suspected manipulations. Moreover, the timeline of each significant bubble is traced back to its emergence time to determine whether the price was influenced by any price-influencing factors demonstrated in the SMD analysis. This allows us to perform an in-depth analysis of market manipulation.

5.2 Approach

GSADF testing is used to evaluate the emergence and duration of bubbles based on historical data regarding currency prices [56]. GSADF tests involve taking all possible sub-samples from a sample dataset (traversing the data starting and ending points), calculating all possible ADF statistics with the sample data, and comparing the maximum GSADF statistic with a 95% critical value on the right. Bubbles are identified when the critical value exceeds the rejection domain. GSADF test results are presented in the form of a table that compares the GSADF t-statistic value with critical values under different confidence levels (α). If the GSADF t-statistic value exceeds critical values, this indicates a bubble in the sampling time of the cryptocurrency. The confidence level α represents the acceptance region for test results in the sample population, and the higher α is, the wider the margin of error is accepted.

According to the methodology used by Phillips et al. [56] in 2015 to calculate the effective window size, the window size is equal to 5% of the total sample size. Minimum effective window X refers to the period within which the bubble value can be calculated. This implies that no statistical test results are available for the first X days of the sample. For example, Dogecoin's effective window size is 74 days, and its sample size is 1,470. Next, we set up a Monte Carlo simulation for 1,000 iterations to ensure sufficient accuracy, and obtained the corresponding statistical value of the GSADF test and corresponding critical value through the 1,000 iterations.

In this work, we first conducted bubble tests for confidence intervals of 90%, 95%, and 99%, respectively, to ensure completeness and rationality, and selected a 95% acceptance region for estimating bubble emergence and duration further. We then calculated a bubble test graph at 95% confidence levels, consisting of a price curve, a critical value curve, and a GSADF t-statistic value curve. In addition to testing the existence of bubbles, this can further determine the time of the

Table 6. GSADF Statistical Test Results

		BTC	ETH	DOGE
GSADF t-Statistic		9.264672	10.88541	11.96402
Test critical values	$\alpha = 99\%$	2.944194	2.944194	2.225943
	$\alpha = 95\%$	2.469119	2.469119	2.500415
	$\alpha = 90\%$	2.21219	2.21219	2.957507

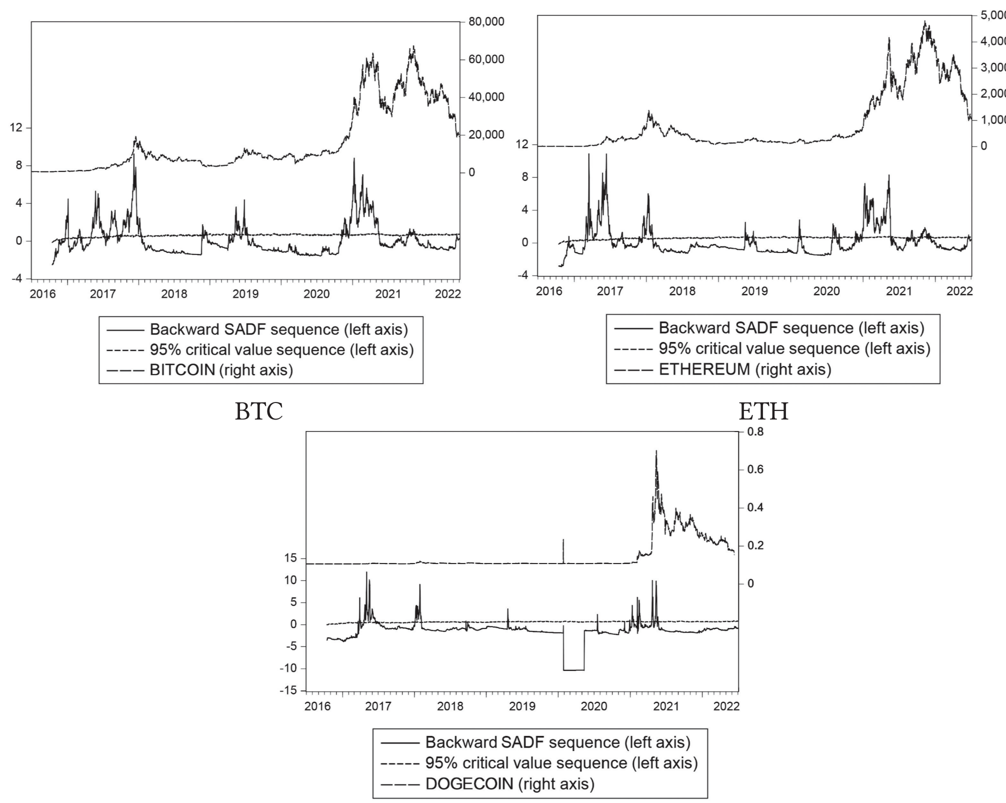

Fig. 5. GSADF test result for three representative cryptocurrencies at a 95% confidence level.

existence of bubbles. When the t-statistic value is greater than the critical value, it means that the price bubble appears at this time point.

5.3 Results

5.3.1 Bubble Existence Analysis. Table 6 presents the results of the bubble existence test. It can be seen that the critical value of each currency at 90%, 95%, and 99% confidence levels is much smaller than the GSADF t-statistic, respectively, indicating that there are significant price bubbles during the sampling period.

5.3.2 Bubble Period Analysis. We examined the periods of time during which bubbles occurred, conducted transactions and SMD, and further examined market manipulation. Figure 5 shows detailed daily GSADF test results for three representative cryptocurrencies.

Table 7. Main Periods of Bitcoin Bubbles

Bubble Periods	Landmark Events
2016.12.21–2017.01.05	Trump's election win
2017.02.23–2017.03.07 2017.05.02–2017.07.09 2017.07.20–2017.09.13	Blockchain commercial applications, Bitcoin ETF and investor enthusiasm
2017.09.27–2018.01.15	Bitcoin futures launch at CME
2018.11.20–2018.12.16	BCH fork and fewer policy restrictions
2019.04.07–2019.07.01	Policy support and market acceptance
2020.11.05–2021.05.14	Global economic stimulus and market speculation, such as Tesla accepting BTC
2021.10.15–2021.11.15	U.S. ETF launch

BTC Analysis. We identified a total of nine major Bitcoin price bubbles, as shown in Table 7. Within these bubble periods, we selected representative SMD and summarized the key reasons or representative events leading to the bubble in the "Landmark Events" column. The largest bubble occurred in 2017, lasting 258 days. Over the next 3 years, the frequency and magnitude of bubbles significantly decreased until the second bubble peaked and the price soared in 2021.

In 2017, the first ICO allowed cryptocurrency founders to sell their new coins directly to investors, leading to an unprecedented wave of speculative activities. An ICO is a concept derived from the concept of IPO in the stock market, and it is the initial issuance of tokens for blockchain projects to raise universal cryptocurrencies such as Bitcoin and Ethereum [49]. In December 2017, the CME (Chicago Mercantile Exchange) officially listed Bitcoin futures, which led to a further increase in the price of Bitcoin, which peaked at $19,511 on December 17, 2017.

The prevailing mentality of "FOMO" (fear of missing the opportunity) [5] and various speculations contributed to the soaring price of Bitcoin, which led to the largest price bubble period in its history. In actuality, the feverish investment sentiment for Bitcoin is more driven by the fear of missing out on appreciation than intrinsic value, and the mining process wastes a tremendous amount of energy, which prompted the Chinese government to require a full liquidation of domestic cryptocurrencies. Subsequently, the United States, Japan, and other countries have become stricter on the regulation of Bitcoin. And with the emergence of Bitcoin's hard fork, BCH, the price of Bitcoin began to plummet, causing the Bitcoin bubble to burst.

Based on our analysis of these price bubble period transactions, we found some suspicious account activities. For example, there was a whale account created in June 2017 with a few transactions. However, each transaction was accompanied by an apparent price fluctuation in Bitcoin. Additionally, the active period of this address is highly correlated with the GSADF test curve. Besides, it appears to be associated with the BitMEX exchange, which has an excellent record for shorting profits throughout its trading history. As an example, on May 17, 2018, the lowest price of Bitcoin on BitMEX was $6,380, whereas the lowest price on Binance, Houbi, and other major exchanges was approximately $7,000. According to Alexander and Heck [2] in 2020, BitMEX clearly manipulated the market by using robots. As of 2020, BitMEX and its founders were charged with financial crimes by the CFTC. Furthermore, BitMEX had been found to have engaged in several unusual transactions with Bitfinex, another exchange that had been accused of manipulating the Bitcoin market during the Bitcoin bubble. In a study by Griffin and Shams [34] published in 2020, it was pointed out that it was mainly Bitfinex that manipulated the Bitcoin boom of 2017 by manipulating USDT, the dollar's token.

ETH Analysis. Table 8 illustrates the timing and key events of six major ETH bubbles. The first and third bubbles lasted the longest at 189 and 186 days, respectively.

Table 8. Main Periods of ETH Bubbles

Bubble Periods	Landmark Events
2017.02.26–2017.07.06	Popularization of ICO and EEA
2017.12.11–2018.02.01	Ethereum switched from PoW to PoS
2019.05.15–2019.06.26	Technological and application progress
2020.02.11–2020.02.23	Government and financial giants are embracing the Ethereum industry
2020.07.30–2020.09.01	Ethereum 2.0
2020.11.21–2021.05.26	The DeFi, NFT hot
2021.8.13–2021.9.6	Ethereum Improvement Proposal 1559
2021.10.15–2021.12.8	Important update to Ethereum-Altair

After the peak of the ETH bubble was reached in 2017, the price of ETH fluctuated approximately in line with its intrinsic value changes over the next 3 years, thus avoiding a concentrated and large-scale bubble. Through 2021, the ETH bubble reached its second peak, with the price also reaching a historical high. Following the establishment of the EEA (the Enterprise Ethereum Alliance) in 2017, the first bubble peak was triggered, reflecting expectations of Ethereum's future potential in the market. From then on, every time a large technology company announced its participation in the EEA, such as Cisco and Bloq, the price of ETH went up. It is important to recognize that the main factor behind the 2017 Ethereum bubble's expansion and collapse was the explosion of ICO projects and market speculative sentiment, which caused the price to deviate from the practical value and technical level of Ethereum at the time. Due to Ethereum's smart contract capabilities that support the creation and development of various applications, more ICOs are being launched and raise ETH based on Ethereum [28]. Many ICO sponsors sold large amounts of the coins they raised to exchange for dollars or other assets quickly, but without actually completing any actual work. Vitalik Buterin remarked the following in [77] with regard to the ICO boom: "A lot of projects will fail and people will lose money." In 2021, ETH experienced the second highest bubble period and the highest and longest uptrend in its history, largely as a result of the popularity of DeFi [63] and NFT. Nevertheless, the frenzied capital chase has highlighted the risks and security issues associated with DeFi systems, which must still be developed and improved.

DOGE Analysis. Unlike ETH and Bitcoin, Dogecoin's price bubbles are quite fierce and can burst very quickly in a very short time. In early 2021, a variety of Dogecoin price bubbles were created and burst in rapid succession, with almost no gap between them. Therefore, we selected 11 major bubble periods from more scattered bubble periods to present in Table 9.

There were three price bubbles for Dogecoin before 2019 but only a slight increase in 2017. The main reason for this is the cryptocurrency boom in 2017, in which Dogecoin had not received significant market attention in terms of trading volume and related news. However, since 2019, every bubble has appeared at the same time as Elon Musk's support for Dogecoin on X. After Elon Musk's initial statement regarding Dogecoin appeared on April 1, 2019, that "Dogecoin might be my favorite cryptocurrency," a bubble rapidly developed. The bubble lasted for only 5 days, but it drove the price of Dogecoin to almost $1. In January 2021, Elon Musk again used X to hype Dogecoin, and more and more supporters came on board. Tesla's acceptance of Dogecoin payments also triggered a speculative flurry. Even though Dogecoin has no intrinsic value, it is closely associated with Elon Musk and Tesla. The price of Dogecoin is even affected by Tesla's stock price. On April 1, 2021, Musk announced his intention to "send Dogecoin to the moon," causing the price of Dogecoin to skyrocket.

Table 9. Main Periods of Dogecoin Bubbles

Bubble Periods	Landmark Events
2017.03.25–2017.06.23	Elon Musk's tweet
2017.12.12–2018.01.14	Cryptocurrency market's boom
2018.09.01–2018.09.10	Dogetherum launch
2019.04.03–2019.04.07	Elon Musk's tweet
2020.07.08–2020.07.08	TikTok Dogecoin frenzy
2020.11.24–2020.11.24	Elon Musk became the world's second richest person
2020.12.20–2020.12.21	Elon Musk's tweet
2021.01.01–2021.01.10	Elon Musk's tweet
2021.01.28–2021.02.14	Elon Musk's tweet
2021.04.13–2021.04.21	Elon Musk's tweet
2021.04.30–2021.05.14	Elon Musk's tweet

We analyzed transactions during these bubble periods and found some addresses that were more active during the price bubble. A few addresses not only coincided with the bubble time but more precisely coincided with the time when Elon Musk tweeted. These addresses were likely to belong to Elon Musk or get inside information about coming tweets. These suspicious addresses are listed in our GitHub repository.

5.4 Findings

Based on the preceding analysis, we summarize the following findings:

— *Finding 7*: Bitcoin bubbles are closely related to the market's acceptance and good news, which makes them susceptible to collapse under restrictions. With tighter policies and expanding market value, it is more difficult to manipulate Bitcoin and the bubble may become more difficult to artificially promote.

— *Finding 8*: The most common method for manipulating the ETH market is boasting about innovation to attract overvalued investments.

— *Finding 9*: Dogecoin is the most vulnerable to direct human influence. Public opinion has directly contributed to the formation of several bubbles, among which Elon Musk's tweets are highly correlated to the price.

6 THREATS TO VALIDITY

6.1 Internal Validity

Manual Efforts. Manual efforts were taken an important role in this article. For example, in Section 3, we manually searched the SMD that resulted in the price fluctuation. Then, we used open card sorting to classify these SMD, which also involved manual efforts. However, we might miss some SMD, and errors might contain in the card sorting process. To reduce errors, all manual efforts were double-checked by two experienced blockchain researchers and any differences were discussed to ensure correctness.

Data Selection. In this article, all analysis was based on a price fall or rise larger than 10%. The threshold of large price fluctuation determines the scale of data we analyzed in this work. Specifically, there might be different market manipulation patterns for price fluctuations lower than 10%, which leads to different findings. Fortunately, we still obtained 75, 159, and 235 price fluctuation days for Bitcoin, Ethereum, and Dogecoin, respectively, based on the current threshold (10%). The number of price fluctuation days is already a statistically representative sample [76] of the whole population, which could prove the integrity of our findings.

Ex-Post Selection. The experimental design used for whale account identification could potentially introduce some degree of ex-post selection bias. Specifically, screening accounts based on pre-determined thresholds post data collection raises the possibility that the results may lack full objectivity or be distorted. This approach also risks failing to comprehensively reflect all accounts warranting suspicion of manipulation. As a consequence, there is a non-trivial possibility that the method overlooks manipulative accounts or inadequately captures the scope of potential market intervention.

6.2 External Validity

Causality vs Correlation. SMD were used to find the reasons leading to price fluctuations. However, it is possible that some SMD were written by unprofessional reporters who misunderstood the reasons for the price fluctuations. To reduce the influence of this situation, all SMD we selected were reported by several professional institutions or individuals.

7 RELATED WORK

Vasek and Moore [68] published the first empirical study on Bitcoin manipulation and scams. By analyzing the reports collected from online forums, they found 192 scams and divided them into four groups: Ponzi schemes, mining scams, scam wallets, and fraudulent exchanges. This work provides evidence that the scam activities could affect Bitcoin price. Gandal et al. [30] conducted an in-depth investigation of price manipulation activities in the Bitcoin ecosystem by using statistical analysis and regression modeling. Chen et al. [17] analyzed the leaked transaction history of Mt. Gox Bitcoin exchanges by using singular value decomposition and found that there was serious market manipulation in the Mt. Gox exchange leading to Bitcoin soaring and dropping several times between 2011 and 2013.

Moore et al. [54] employed a logistic regression model to study the phenomenon of large-scale fraud in the Bitcoin market between 2010 and 2015. Based on correlation analysis of transaction statistics, Wei [70] concluded that significant amounts of capital were used to manipulate the price of Bitcoin by Tether. Additionally, to analyze market manipulation from transaction data, Mirtaheri et al. [53] investigated cryptocurrency market manipulation from the perspective of social media. Fan et al. [27] proposed a method for detecting smart contract scams by mining the topological features of transaction data. They extracted interaction features from dynamic transaction information and analyzed interaction features and topological structure data, enabling discovery of fraudulent smart contracts on Ethereum. In recent years, other cryptocurrencies have also been the subject of market manipulation scandals. Chen et al. [19] expand the research on detecting the Ponzi scheme on Ethereum based on previous research. Chen et al. [15] further conducted another study specifically focusing on phishing scams in the cryptocurrency field. Xia et al. [74] conducted a horizontal analysis of fraudulent tokens on Uniswap, determining that 50% of scam tokens were specifically set up for "pull and rug" schemes, with some scam tokens planting backdoors and vulnerabilities in their smart contracts. With in-depth studies of crypto scams, much attention has been paid to the "pump and dump" manipulation of cryptocurrencies, which involves using speculation and other methods to raise prices and dumping quickly [18, 38, 45]. La Morgia et al. [46] leverage a unique dataset of the verified pump and dumps to build a machine learning model able to detect a pump and dump scam.

Many previous works believe that speculative manipulation is primarily responsible for the emergence of bubbles in the cryptocurrency market [4, 20, 44]. Many studies have indicated that the Bitcoin price bubble, which began in June 2015, was accompanied by speculative manipulation and unnatural market behavior [12]. However, detecting price bubbles can be achieved by utilizing a variety of approaches. By using the unit root test method, Malhotra and Maloo [51] demonstrated that the Bitcoin market is experiencing a price bubble. Cheung et al. [20] used a robust bubble detection method to prove the existence of speculative bubbles. According to the rapid growth of cryptocurrency market, researchers have investigated other cryptocurrencies for signs of bubble characteristics. Corbet et al. [23] further examined the existence and dates of pricing bubbles in Bitcoin and Ethereum, and Fry and Cheah [29] extended their conclusions to the broader cryptocurrency market. In the study conducted by de Souza et al. [24], all cryptocurrencies, including ETH and Ripple, experienced bubbles over the past few years.

In terms of blockchain data analysis, Chen et al. [16] collected all of the transactions on Ethereum and leveraged graph analysis to conduct the first systematic study on the Ethereum ecosystem. Their work gives a blueprint of Ethereum money transfer, smart contract creation, and invocation. McGinn et al. [52] proposed a method to visualize Bitcoin transactions to explore users' activities. Their work found that money laundering and attack activities were popular on the Bitcoin market.

Most price bubble analyses focus on the stock market and have a long history. In 1981, LeRoy and Porter [47] proposed the first quantitative analysis method for determining whether bubbles exist by observing variance changes, but this method was too simple and inaccurate. Engle and Granger [26] proposed a co-integration test method, but this method is more suitable for currency fluctuations with linear evolution. Phillips et al. [57] proposed the *Sup ADF* test method with higher accuracy by combining forward recursive regression and Dickey-Fuller right unit root tests. However, the Sup ADF test was not effective enough for the data with multiple bubbles. Inspired by the Sup ADF model, in a different work, Phillips et al. [56] improved the testing model and proposed new methods named *GSADF*, which was used in this work to detect the bubbles of the cryptocurrency market.

In terms of methodology, many empirical studies on cryptocurrencies have also utilized methods such as data engineering, NLP, and modeling analysis. For instance, Hayes [35] analyzed cross-sectional data from 66 commonly used cryptocurrencies, identifying the key driving factors of cryptocurrency value. Dhawan and Putniņš [25] examined data from 355 trading pairs over 6 months, uncovering a significant wealth transfer between abnormal trading volumes and whales, providing strong empirical evidence for extreme price distortion. Chen et al. [13] conducted empirical research by collecting off-chain data from five exchanges websites and on-chain data.

In this work, we conduct three analyses to investigate the market manipulation of cryptocurrencies: SMD analysis, blockchain data analysis, and price bubble analysis. To the best of our knowledge, the current SMD analysis mainly focuses on finding scams. For example, Vasek and Moore [68] published the first empirical study on Bitcoin-based scams. By analyzing the reports collected from online forums, they found 192 scams and divided them into four groups. Chen et al. [14] analyzed the online posts on a Q&A website (Stack Exchange). They used open card sorting to summarize 20 Ethereum smart contract defects and divided them into five groups.

To conclude, we combine previous analysis into a comprehensive discussion of market manipulation with no restrictions on trading platforms and partial information websites from two perspectives: transaction and social media. Additionally, most previous studies focused primarily on Bitcoin rather than the entire cryptocurrency market, but in the past few years, with the boom of cryptocurrencies and the improvement of the Bitcoin regulatory mechanism, cryptocurrency has become the new favorite of investors and the protagonist of a variety of speculative cases. As the cryptocurrency market is rapidly evolving, distant data may not be sufficient for current reference; we intend to eliminate the time and data source limitations by developing new platforms and tools.

8 CONCLUSION AND FUTURE WORK

We conducted a comprehensive analysis of market manipulation of cryptocurrencies by using a combination of SMD analysis, blockchain data analysis, and price bubble analysis. Using SMD analysis, we summarized seven kinds of SMD that lead to price fluctuation in the cryptocurrency market. Bitcoin price has a high correlation to government policy, especially for China, and individuals have little responsibility for the price fluctuation of Bitcoin. The technology factor plays an important role in Ethereum price fluctuation. Market manipulation is prevalent on Dogecoin. Elon Musk has a strong influence on Dogecoin's price. Our blockchain transaction data analysis shows that Ethereum and Dogecoin have great centralization, which gives whale users the ability to manipulate the price. Bitcoin, Ethereum, and Dogecoin all seem to have some "abnormal accounts" able to "predict" the future cryptocurrency price. These abnormal accounts suggest the existence of some form of insider trading and manipulation of the cryptocurrencies market. Only a limited number of abnormal accounts are found on Bitcoin, which shows that market manipulation is not common in the Bitcoin market. Our price bubble analysis confirmed that there is less market manipulation in Bitcoin. Technical companies/teams usually hype some new concepts (e.g., ICO and DeFi), which tend to drive up the price of Ethereum. Elon Musk's tweets are highly correlated with the price of Dogecoin. Based on the manipulation patterns and abnormal accounts that we found, we plan to design a system to detect market manipulation and warn of the risk for investors when manipulation happens.

In our future work, we plan to employ additional quantitative analysis techniques to strengthen our research findings. For instance, we will utilize neural networks to model the non-linear dependencies within price sequences and identify the fluctuation patterns influenced by manipulation. Additionally, we intend to apply methods rooted in behavioral finance theories, such as attention networks, to identify thematic patterns in the trading behavior of manipulators. Furthermore, we will explore the utilization of generative adversarial networks and other methods to synthetically replicate manipulative trading behavior, thereby evaluating its actual impact on the market.

Moreover, we have plans to expand the sample size and time range to conduct empirical analyses and unveil manipulation characteristics across different currency markets. We genuinely appreciate your valuable feedback and suggestions. They will serve as significant references for our future work in enhancing the quality and reliability of our research. We will strive to make improvements accordingly and eagerly anticipate achieving better research outcomes.

REFERENCES

[1] Adebayo Adedokun. 2019. Bitcoin-Altcoin price synchronization hypothesis: Evidence from recent data. *Journal of Finance and Economics* 7, 4 (2019), 137–147.

[2] C. Alexander and D. F. Heck. 2020. Price discovery in Bitcoin: The impact of unregulated markets. *Journal of Financial Stability* 50 (2020), 100776.

[3] Lennart Ante. 2023. How Elon Musk's Twitter activity moves cryptocurrency markets. *Technology Forecasting and Social Change* 186, A (2023), 122112.

[4] Paul Barnes 2018. Crypto currency and its susceptibility to speculative bubbles, manipulation, scams and fraud. *Journal of Advanced Studies in Finance* 9, 2 (18) (2018), 60–77.

[5] Dirk G. Baur and Thomas Dimpfl. 2018. Asymmetric volatility in cryptocurrencies. *Economics Letters* 173 (2018), 148–151.

[6] Walter Bazán-Palomino. 2021. How are Bitcoin forks related to Bitcoin? *Finance Research Letters* 40 (2021), 101723.

[7] David M. Blei and John D. Lafferty. 2007. A correlated topic model of science. *Annals of Applied Statistics* 1, 1 (2007), 17–35.

[8] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent Dirichlet allocation. *Journal of Machine Learning Research* 3 (2003), 993–1022.

[9] Dmitri Boreiko and Navroop K. Sahdev. 2018. To ICO or not to ICO—Empirical analysis of initial coin offerings and token sales. *SSRN*. Retrieved February 5, 2024 from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3209180

[10] Jamal Bouoiyour, Refk Selmi, Aviral Kumar Tiwari, and Olaolu Olayeni. 2016. What drives Bitcoin price. *Economics Bulletin* 36, 2 (2016), 843–850.

[11] Vitalik Buterin. 2013. *Ethereum White Paper*. GitHub Repository.

[12] Eng-Tuck Cheah and John Fry. 2015. Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters* 130 (2015), 32–36.
[13] Jialan Chen, Dan Lin, and Jiajing Wu. 2022. Do cryptocurrency exchanges fake trading volumes? An empirical analysis of wash trading based on data mining. *Physica A: Statistical Mechanics and Its Applications* 586 (2022), 126405.
[14] Jiachi Chen, Xin Xia, David Lo, John Grundy, Xiapu Luo, and Ting Chen. 2022. Defining smart contract defects on Ethereum. *IEEE Transactions on Software Engineering* 48, 1 (2022), 327–345.
[15] Liang Chen, Jiaying Peng, Yang Liu, Jintang Li, Fenfang Xie, and Zibin Zheng. 2020. Phishing scams detection in Ethereum transaction network. *ACM Transactions on Internet Technology* 21, 1 (2020), 1–16.
[16] Ting Chen, Yuxiao Zhu, Zihao Li, Jiachi Chen, Xiaoqi Li, Xiapu Luo, Xiaodong Lin, and Xiaosong Zhange. 2018. Understanding Ethereum via graph analysis. In *Proceedings of the 2018 IEEE Conference on Computer Communications (INFOCOM'18)*. IEEE, 1484–1492.
[17] Weili Chen, Jun Wu, Zibin Zheng, Chuan Chen, and Yuren Zhou. 2019. Market manipulation of Bitcoin: Evidence from mining the Mt. Gox transaction network. In *Proceedings of the 2019 IEEE Conference on Computer Communications (INFOCOM'19)*. 964–972.
[18] Weili Chen, YueJin Xu, Zibin Zheng, Yuren Zhou, Jianxun Eileen Yang, and Jing Bian. 2019. Detecting "pump & dump schemes" on cryptocurrency market using an improved apriori algorithm. In *Proceedings of the 2019 IEEE International Conference on Service-Oriented System Engineering (SOSE'19)*. IEEE, 293–2935.
[19] Weili Chen, Zibin Zheng, Jiahui Cui, Edith Ngai, Peilin Zheng, and Yuren Zhou. 2018. Detecting Ponzi schemes on Ethereum: Towards healthier blockchain technology. In *Proceedings of the 2018 World Wide Web Conference*. 1409–1418.
[20] Adrian Cheung, Eduardo Roca, and Jen-Je Su. 2015. Crypto-currency bubbles: An application of the Phillips–Shi–Yu (2013) methodology on Mt. Gox Bitcoin prices. *Applied Economics* 47, 23 (2015), 2348–2358.
[21] Usman W. Chohan. 2021. A History of Dogecoin. Discussion Series: Notes on the 21st Century. Retrieved February 5, 2024 from https://ssrn.com/abstract=3091219
[22] Pavel Ciaian, Miroslava Rajcaniova, and d'Artis Kancs. 2018. Virtual relationships: Short- and long-run evidence from BitCoin and Altcoin markets. *Journal of International Financial Markets, Institutions and Money* 52 (2018), 173–195.
[23] Shaen Corbet, Brian Lucey, and Larisa Yarovaya. 2018. Datestamping the Bitcoin and Ethereum bubbles. *Finance Research Letters* 26 (2018), 81–88.
[24] Michel Cândido de Souza, Elder Tiago da Costa de Souza, and Hugo Carcanholo Iasco Pereira. 2017. Cryptocurrencies bubbles: New evidences. *Empirical Economics Letters* 16, 7 (2017), 739–746.
[25] Anirudh Dhawan and Tālis J. Putniņš. 2023. A new wolf in town? Pump-and-dump manipulation in cryptocurrency markets. *Review of Finance* 27, 3 (2023), 935–975.
[26] R. F. Engle and C. W. J. Granger. 1987. Co-integration and error-correction: Representation, estimation and testing. *Econometrica* 55, 2 (1987), 251–276.
[27] Shuhui Fan, Shaojing Fu, Yuchuan Luo, Haoran Xu, Xuyun Zhang, and Ming Xu. 2022. Smart contract scams detection with topological data analysis on account interaction. In *Proceedings of the 31st ACM International Conference on Information and Knowledge Management*. 468–477.
[28] Gianni Fenu, Lodovica Marchesi, Michele Marchesi, and Roberto Tonelli. 2018. The ICO phenomenon and its relationships with Ethereum smart contract environment. In *Proceedings of the 2018 International Workshop on Blockchain Oriented Software Engineering (IWBOSE'18)*. IEEE, 26–32.
[29] John Fry and Eng-Tuck Cheah. 2016. Negative bubbles and shocks in cryptocurrency markets. *International Review of Financial Analysis* 47 (2016), 343–352.
[30] Neil Gandal, J. T. Hamrick, Tyler Moore, and Tali Oberman. 2018. Price manipulation in the Bitcoin ecosystem. *Journal of Monetary Economics* 95 (2018), 86–96.
[31] David Garcia, Claudio J. Tessone, Pavlin Mavrodiev, and Nicolas Perony. 2014. The digital traces of bubbles: Feedback cycles between socio-economic signals in the Bitcoin economy. *Journal of the Royal Society Interface* 11, 99 (2014), 20140623.
[32] Adem Efe Gencer, Soumya Basu, Ittay Eyal, Robbert van Renesse, and Emin Gün Sirer. 2018. Decentralization in Bitcoin and Ethereum networks. In *Proceedings of the International Conference on Financial Cryptography and Data Security*. 439–457.
[33] Google. 2021. Google BigQuery Introduction. Retrieved Dec 28, 2021 from https://cloud.google.com/bigquery
[34] John M. Griffin and Amin Shams. 2020. Is Bitcoin really untethered? *Journal of Finance* 75, 4 (2020), 1913–1964.
[35] Adam Hayes. 2015. What factors give cryptocurrencies their value: An empirical analysis. *SSRN*. Retrieved February 5, 2024 from https://ssrn.com/abstract=2579445
[36] Adam S. Hayes. 2019. Bitcoin price and its marginal cost of production: Support for a fundamental value. *Applied Economics Letters* 26, 7 (2019), 554–560.

[37] Frankenfield Jake and Overcast Kimberly. 2022. Initial Coin Offering (ICO): Coin Launch Defined, with Examples. Retrieved January 3, 2022 from https://www.investopedia.com/terms/i/initial-coin-offering-ico.asp
[38] Josh Kamps and Bennett Kleinberg. 2018. To the moon: Defining and detecting cryptocurrency pump-and-dumps. *Crime Science* 7, 1 (2018), 1–18.
[39] Ladislav Kristoufek. 2015. What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. *PLoS ONE* 10, 4 (2015), e0123923.
[40] Ladislav Kristoufek. 2019. Is the Bitcoin price dynamics economically reasonable? Evidence from fundamental laws. *Physica A: Statistical Mechanics and Its Applications* 536 (2019), 120873.
[41] Ladislav Kristoufek. 2023. Will Bitcoin ever become less volatile? *Finance Research Letters* 51 (2023), 103353.
[42] Jan Kubal and Ladislav Kristoufek. 2022. Exploring the relationship between Bitcoin price and network's hashrate within endogenous system. *International Review of Financial Analysis* 84 (2022), 102375. https://doi.org/10.1016/j.irfa.2022.102375
[43] Sinan Küfeoğlu and Mahmut Özkuran. 2019. Bitcoin mining: A global review of energy and power demand. *Energy Research & Social Science* 58 (2019), 101273.
[44] Nikolaos Kyriazis, Stephanos Papadamou, and Shaen Corbet. 2020. A systematic review of the bubble dynamics of cryptocurrency prices. *Research in International Business and Finance* 54 (2020), 101254.
[45] Massimo La Morgia, Alessandro Mei, Francesco Sassi, and Julinda Stefa. 2020. Pump and dumps in the Bitcoin era: Real time detection of cryptocurrency market manipulations. In *Proceedings of the 2020 29th International Conference on Computer Communications and Networks (ICCCN'20)*. IEEE, 1–9.
[46] Massimo La Morgia, Alessandro Mei, Francesco Sassi, and Julinda Stefa. 2021. The Doge of wall street: Analysis and detection of pump and dump cryptocurrency manipulations. *arXiv preprint arXiv:2105.00733* (2021).
[47] S. F. Leroy and R. D. Porter. 1981. Stock price volatility: A test based on implied variance bounds. *Econometrica* 49, 3 (1981), 555–574.
[48] Xin Li and Chong Alex Wang. 2017. The technology and economic determinants of cryptocurrency exchange rates: The case of Bitcoin. *Decision Support Systems* 95 (2017), 49–60. https://doi.org/10.1016/j.dss.2016.12.001
[49] Daniel Liebau and Patrick Scheuffel. 2019. Cryptocurrencies & initial coin offerings: Are they scams? An empirical study. *Journal of British Blockchain Association* 2, 1 (2019), 1–7.
[50] Lukáš Pichl and Taisei Kaizoji. 2017. Volatility analysis of Bitcoin price time series. *Quantitative Finance and Economics* 1 (2017), 474–485.
[51] Akash Malhotra and Mayank Maloo. 2014. Bitcoin—Is it a bubble? Evidence from unit root tests. In *Proceedings of the 2014 Conference on Evidence from Unit Root Tests.*
[52] Dan McGinn, David Birch, David Akroyd, Miguel Molina-Solana, Yike Guo, and William J. Knottenbelt. 2016. Visualizing dynamic Bitcoin transaction patterns. *Big Data* 4, 2 (2016), 109–119.
[53] Mehrnoosh Mirtaheri, Sami Abu-El-Haija, Fred Morstatter, Greg Ver Steeg, and Aram Galstyan. 2021. Identifying and analyzing cryptocurrency manipulations in social media. *IEEE Transactions on Computational Social Systems* 8, 3 (2021), 607–617.
[54] Tyler Moore, Nicolas Christin, and Janos Szurdi. 2018. Revisiting the risks of Bitcoin currency exchange closure. *ACM Transactions on Internet Technology* 18, 4 (2018), 1–18.
[55] Satoshi Nakamoto. 2008. A peer-to-peer electronic cash system. Retrieved February 5, 2024 from https://bitcoin.org/bitcoin.pdf
[56] P. C. B Phillips, S. P. Shi, and J. Yu. 2013. *Testing for Multiple Bubbles 1: Historical Episodes of Exuberance and Collapse in the S&P 500*. School of Economics, Singapore Management University.
[57] Peter C. B. Phillips, Yangru Wu, and Jun Yu. 2011. Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review* 52, 1 (2011), 201–226.
[58] Marc Pilkington. 2016. Blockchain technology: Principles and applications. In *Research Handbook on Digital Transformations*. Edward Elgar Publishing.
[59] Burak Pirgaip, Burcu Dinçergök, and Şüheda Haşlak. 2019. Bitcoin market price analysis and an empirical comparison with main currencies, commodities, securities and Altcoins. In *Blockchain Economics and Financial Market Innovation: Financial Innovations in the Digital Age*. Springer, 141–166.
[60] Rakesh Sharma and Amilcar Chavarria. 2022. What Is Decentralized Finance (DeFi) and How Does It Work? Retrieved January 13, 2022 from https://www.investopedia.com/decentralized-finance-defi-5113835
[61] Raynor de Best. 2021. Cryptocurrency Market Value. Retrieved December 29, 2021 from https://www.statista.com/statistics/730876/cryptocurrency-maket-value/
[62] Svetlana Sapuric and Angelika Kokkinaki. 2014. Bitcoin is volatile! Isn't that right? In *Business Information Workshops*. Lecture Notes in Business Information Processing, Vol. 183. Springer, 255–265.
[63] Fabian Schär. 2021. Decentralized finance: On blockchain-and smart contract-based financial markets. *SSRN*. Retrieved February 5, 2024 from https://ssrn.com/abstract=3571335

[64] Yahya Shahsavari, Kaiwen Zhang, and Chamseddine Talhi. 2019. A theoretical model for fork analysis in the Bitcoin network. In *Proceedings of the 2019 IEEE International Conference on Blockchain (Blockchain'19)*. IEEE, 237–244.

[65] Syed Jawad Hussain Shahzad, Muhammad Anas, and Elie Bouri. 2022. Price explosiveness in cryptocurrencies and Elon Musk's Tweets. *Finance Research Letters* 47, B (2022), 102695.

[66] Yhlas Sovbetov. 2018. Factors influencing cryptocurrency prices: Evidence from Bitcoin, Ethereum, Dash, Litcoin, and Monero. *Journal of Economics and Financial Analysis* 2, 2 (2018), 1–27.

[67] Donna Spencer. 2009. *Card Sorting: Designing Usable Categories*. Rosenfeld Media.

[68] Marie Vasek and Tyler Moore. 2015. There's no free lunch, even using Bitcoin: Tracking the popularity and profits of virtual currency scams. In *Proceedings of the International Conference on Financial Cryptography and Data Security*. 44–61.

[69] Dejan Vujičić, Dijana Jagodić, and Siniša Ranđić. 2018. Blockchain technology, Bitcoin, and Ethereum: A brief overview. In *Proceedings of the 2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH'18)*. IEEE, 1–6.

[70] Wang Chun Wei. 2018. The impact of tether grants on Bitcoin. *Economics Letters* 171 (2018), 19–22.

[71] Wikipedia. 2021. Dogecoin. Retrieved December 29, 2021 from https://en.wikipedia.org/wiki/Dogecoin

[72] Ines Wöckl. 2019. Bubble detection in financial markets—A survey of theoretical bubble models and empirical bubble detection tests. *SSRN*. Retrieved February 5, 2024 from https://ssrn.com/abstract=3460430

[73] Gavin Wood. 2017. *Ethereum: A Secure Decentralised Generalised Transaction Ledger*. Yellow Paper. Ethereum.

[74] Pengcheng Xia, Haoyu Wang, Bingyu Gao, Weihang Su, Zhou Yu, Xiapu Luo, Chao Zhang, Xusheng Xiao, and Guoai Xu. 2021. Trade or trick? Detecting and characterizing scam tokens on Uniswap decentralized exchange. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 5, 3 (2021), 1–26.

[75] Yuan Yiming, Du Hai, Shi Junjing, and Xiao Xiao. 2021. Huobi Weekly Report Analysis. Vol. 26. Retrieved Aug 17, 2018, from https://medium.com/hbus-official/big-data-weekly-26-cbc20cd979c9

[76] Julie Young. 2021. Representative Sample. Retrieved October 29, 2021 from https://www.investopedia.com/terms/r/representative-sample.asp

[77] Avi Mizrahi. 2017. Ethereum founder vitalik buterin: We are in an ICO bubble. Finance Magnates. https://www.financemagnates.com/cryptocurrency/news/ethereum-founder-vitalik-buterin-ico-bubble/

Received 30 January 2023; revised 21 November 2023; accepted 25 January 2024