Changes in the market structure and risk management of Bitcoin and its forked coins

Xiaolin Kong
Business School, Hunan University, China
Research Institute of Digital Society and Blockchain, Hunan University, China
Email: kongxiaolin@hnu.edu.cn

Chaoqun Ma
Business School, Hunan University, China
Research Institute of Digital Society and Blockchain, Hunan University, China
Center for Resource and Environmental Management, Hunan University, China
Email: cqma_hnu@163.com

Yi-Shuai Ren*
School of Public Administration, Hunan University, China
Research Institute of Digital Society and Blockchain, Hunan University, China
Center for Resource and Environmental Management, Hunan University, China
The Energy Centre, University of Auckland, 12 Grafton Rd, Auckland, 1010, New Zealand
Email: renyishuai1989@126.com

Seema Narayan
Asia Pacific Applied Economics Association, Melbourne, Australia
Email: swdhar27@gmail.com

Thong Trung Nguyen
School of Banking, University of Economics Ho Chi Minh City, Vietnam
Email: thongnt@ueh.edu.vn

Konstantinos Baltas
Essex Business School, University of Essex, United Kingdom
Email: k.baltas@essex.ac.uk

Corresponding author email: renyishuai1989@126.com

* Corresponding author.
E-mail: renyishuai1989@126.com (Yi-Shuai Ren).

Changes in the market structure and risk management of Bitcoin and its forked coins

Abstract

Inconsistency of consensus results in blockchain forks, which create a new financial risk. After filtering out Bitcoin's linear, nonlinear, and lag impacts on forked coins, this study employs a bottom-up hierarchical clustering algorithm to examine the logarithmic return series for Bitcoin and its 14 forked coins from 2018 to 2021. The results indicate that the market for forked coins can be divided into three clusters: SegWit-supported forked coins, mature forked coins, and the latest forked coins. Bitcoin and the mature forked coins form a cluster, and its performance is superior to others. Although Bitcoin's return significantly affects that of its forked coins, it does not affect the market structure. Furthermore, this study provides references for risk aversion among investors in forked coins and presents macro-level information for cryptocurrency market authorities.

Keywords: Bitcoin; Forked coin; Market structure; Financial risk; Hierarchical clustering algorithm; Blockchain

1. Introduction

Nakamoto (2008) first described the blockchain concept. This technology can be viewed as a decentralized distributed ledger, with nodes sharing the blockchain's specific content (Iansiti and Lakhani, 2017). Bitcoin (BTC), among the most well-known blockchain applications, was created amid the 2008 global financial crisis. BTC utilizes proof of work (PoW) to build consensus. In the BTC blockchain, miners can receive BTCs if they obtain accounting rights. Additionally, the BTC blockchain has certain glaring flaws, notably a block size of only 1Mb, resulting in low work efficiency and an inability to record many transactions on time. Furthermore, BTC blockchain users risk losing their private keys (Van Alstyne, 2014) and high transaction fees (Poon and Dryja, 2016). Therefore, block expansion has become a critical issue that must be immediately resolved; however, a blockchain will fork if the community cannot reach a consensus on the upgrade plan (Islam et al., 2019).

Forking behavior is a change in a blockchain's rules that forks it into two or more potential chains (Nyman et al., 2012). Specifically, two different forms of the forking phenomenon exist: soft and hard forks (see Figure 1). After the blockchain has been improved, nonupgraded nodes can still accept the information on the chain; this is a soft fork. In contrast, a hard fork requires all nodes to be upgraded; Bitcoin Cash (BCH) is the most prominent example. According to the statistics of the forked coin information website, *forkdrop.io*, there are 105 BTC-forked coins thus far. BTC-forked coins have endlessly emerged in the cryptocurrency market (Bowden, 2018)[1].

<Insert Figure 1 Here>

The regular occurrence of blockchain forks, not only on the BTC blockchain (Tomić, 2020), can result in numerous issues. First, the fork will incur tax and legal risks (Webb, 2018; Chason, 2019b; Cotler, 2020), followed by astronomical research and development (R&D) expenditures (Schär, 2020). Second, forked coins are similar, typically for marketing goals rather than for practical reasons. Third, an unsuccessful fork will cause significant price changes in BTC (Chaim and Laurini, 2018). Due to the

---

[1] The reference basis for the statement on forked coins is https://support.exodus.com/article/203-how-can-i-claim-my-forked-coins.

different versions of the two ledgers, the blockchain's credibility and reliability will be diminished (Biais et al., 2019). The frequent occurrence of forks and various corresponding risks increase the need to conduct related research on forks and forked coins. With blockchain development, the market for forked coins has only begun. Nonetheless, it should also garner sufficient regulatory attention to be included in the macro-prudential regulatory framework to prevent systemic risks.

Numerous studies have examined forks and forked coins (Bowden, 2018; Webb, 2018; Misic et al., 2019; Cotler, 2020; Kubiak and Kutylowski, 2020; Schär, 2020; Tomić, 2020; Bazán-Palomino, 2021; Johnson, 2021). From a technological standpoint, many studies have investigated the causes of forking behavior (Biais, 2019; Schär, 2020), while from an economic standpoint, some studies focus on the price or revenue relationship between BTC and forked coins (Yi et al., 2021; Bazán-Palomino, 2021). Numerous studies have examined a fork's tax implications from a financial perspective (Xu, 2019; Chason, 2019a; Cotler, 2020; Webb, 2020). Despite considerable research on forked coins, few researchers have investigated their global aspects, such as their market structure.

Consequently, we employ a hierarchical clustering algorithm to examine the structural characteristics of forked coins traded on the market between 2018 and 2021. While BTC significantly affects the return on forked coins, it has no meaningful effect on the cluster. In addition, we classify the coins into three clusters based on their return characteristics.

This study's contributions mainly include the following aspects: (1) Exploring the market structure for forked coins closes gaps in the realm of cryptocurrencies. Our division performs empirical structural verification. (2) This study applies the filtering method to eliminate BTC's effect, and we address BTC's linear, nonlinear, and time-lag impacts. (3) Our research suggests that BTC and its forked coins are interconnected or share comparable properties. As far as we know, although several technical studies have investigated forks, no direct financial research has examined forked coins. (4) Our findings provide reference information for the regulatory authorities and help cryptocurrency investors minimize market risk.

The remainder of this study is organized as follows: Section 2 presents the literature review. In Section 3, we explain the data and methodology. Section 4 provides the results and discussion, while Section 5 concludes and offers policy implications.

2. Literature review

Studies related to our research can be separated into two components. One is the investigation of BTC and its risks, including its financial, technological, and market characteristics. Due to the possible risks of BTC, forked coins have emerged. The other component focuses on forking phenomena, which provide a comprehensive view of the research on forked coins.

2.1 Risk characteristics of BTC

Although BTC is the first and most well-known blockchain application, it has several technical and application-related issues (Hu et al., 2019). Consequently, research on the risks of BTC is ongoing.

Regarding technology, the BTC blockchain wastes resources due to the PoW consensus process, and miners consume significant electrical resources in mining (Fauzi et al., 2020). In addition to resource loss, the blockchain is a decentralized distributed ledger, and the absence of a central supervisor renders it susceptible to various attacks. Courtois and Bahack (2014) summarized many assault types in BTC, including the pool-hopping, mining-cartel, and difficulty-raising attacks.

Regarding finance, the volatility of the BTC price has aroused widespread concern among researchers (Al Mamun et al., 2020), and academics have presented a range of BTC risk-measurement methods. For example, Troster et al. (2019) assessed the performance of 45 models and discovered that heavy-tailed generalized autoregressive score (GAS) might adequately explain the risk features of BTC. With the idea of considering outliers, Trucios (2019) developed a range of robust algorithms for assessing BTC volatility, including robust generalized autoregressive conditional heteroskedasticity (GARCH) and robust GAS algorithms. Li et al. (2019) used the generalized augmented Dickey-Fuller (ADF) algorithm to calculate the period for BTC price bubbles. They discovered that these bubbles primarily occurred during periods of rising BTC prices. Lu et al. (2021) integrated the long- and short-term memory (LSTM) model with the value at risk (VaR) and expected shortfall models. A portfolio-forecasting model was developed to examine BTC's market risk and accounts for cryptocurrency returns' nonlinearity and long-term reliance. The study of the risk posed by BTC has attracted considerable attention.

Moreover, BTC's financial risks extend to other cryptocurrencies and possibly conventional financial markets. In addition to BTC, dozens of other cryptocurrencies are traded on the cryptocurrency market, and academics have extensively examined the relationship between BTC and these other cryptocurrencies. Cagli (2019), for instance, utilized BTC and seven other cryptocurrencies as samples to investigate the explosive behavior of cryptocurrencies, demonstrating a substantial, pairwise, comovement correlation between cryptocurrencies with explosive behavior. Ciaian and Rajcaniova (2018) used the autoregressive distributed lag (ARDL) model to show a considerable dependence between BTC and 16 other cryptocurrencies and that the short-term correlation was stronger than the long-term relationship. Demir et al. (2021) established a nonlinear ARDL model to describe the impact of BTC on the prices of three other cryptocurrencies, confirming the existence of the BTC impact and highlighting that when BTC price declined, the impact on other cryptocurrencies was greater than when BTC price rose.

Inseparable from the traditional financial market is the cryptocurrency market represented by BTC. Dyhrberg (2016) placed BTC between gold and the United States (US) dollar (USD). Kurka (2019) examined the asymmetric transmission of shocks between the cryptocurrency and traditional financial markets, validating the bidirectional transmission of big shocks between the two sectors. Furthermore, Bouri et al. (2020) employed wavelet analysis to demonstrate BTC's poor dependence on gold, commodities, and stocks at various time scales. As the risk of cryptocurrencies may also be passed on to traditional financial markets, conducting an in-depth study on BTC is vital.

2.2 BTC and its forked coins

With the rise of blockchain forks in recent years, forked coins have also generated interest. For example, Bazán-Palomino (2020) utilized three multivariate volatility models to calculate the dynamic correlation between BTC and its forked coins, determining that the time-varying correlation was negative during periods of extreme volatility. In subsequent investigations, Bazán-Palomino (2021) utilized the GARCH model to examine the volatility correlation between BTC and its forked coins, revealing that their interaction was substitutional during market volatility and complimentary during market stability. Consequently, the relationship between BTC and forked coins has been demonstrated.

BCH is among the earliest and most well-known coins with a forked design; thus, it has received significant attention. Johnson (2021) examined the impact of BTC price on BCH by constructing a multiplier indicator and validated the existence of a correlation; except for the first six months of blockchain forking, the price trends for BTC and BCH stayed more than 75% consistent. In addition to studies that examine the correlation between BTC and forked coins, Yi et al. (2021) examined the conversion of information flow utilizing transfer entropy. They discovered that information flowed in the mode of price increases and decreases. These studies examine the relationship between BTC and forked coins and demonstrate that BTC significantly affects forked coins. When examining BTC and forked coins, we cannot overlook the impact of the former's price on that of the latter.

Nonetheless, more forked-coin research focuses on the technological field (Lin and Liao, 2017; Möser et al., 2017; Vujičić et al., 2018; Misic et al., 2019; Kubiak and Kutylowski, 2020; Bazán-Palomino, 2021; Liu et al., 2021). Each block in the BTC blockchain is only 2Mb, limiting the number of transactions that may be supported. Consequently, the amount of transactions that the BTC blockchain can process per unit of time is limited. Insufficient business volume in the first few years of BTC's operation impeded the blockchain's development; however, as the price of BTC, number of executed transactions, and transaction fees skyrocketed in 2017, the inefficiencies of the previous version of the BTC blockchain required immediate attention (Göbel and Krzesinski, 2017). Forking is the sole remedy for the abovementioned issues (Bazán-Palomino, 2020). Islam et al. (2019) defined blockchain forking behavior as "*Changes in a blockchain's rules that lead to permanent divergence of the blockchain and its development into two or more potential paths*." Soft and hard forks in the blockchain were identified as answers to the two techniques of inconsistent blockchain-consensus behavior (Yi et al., 2021).

Currently, no analysis exists for the market's overall characteristics for forked coins; however, research on the structure of the financial market has a high reference value for our study. The clustering algorithm is a common data-mining technique employed in studying financial market structure (Hsu et al., 2000). Researchers divide assets into clusters using hierarchical clustering, k-means, and other techniques based on the similarity of return series. On the one hand, this approach can assist regulators in classifying and supervising, while on the other, it can enhance the diversification of asset allocation and lessen the market risks experienced by investors. For instance, Momeni et al. (2015) classified firms on the Tehran Stock Exchange using the k-means method. According to their profitability, companies in the sample’s three industries were grouped into two clusters.

Furthermore, regulators can establish differentiated oversight for various types of businesses. Aste (2019) calculated the Kendell correlation coefficient and nonparametric transfer entropy between the cryptocurrency and market sentiment and subsequently demonstrated that, according to the structural characteristics of the cryptocurrency market, significant cryptocurrencies, such as BTC, played a central role. Nanda et al. (2010) evaluated the stock-market structure of the Bombay Exchange in India from 2007 to 2008 using three different algorithms, including the k-means algorithm, self-organizing maps (SOMs), and fuzzy c-means, concluding that k-means performed the best. They created a portfolio for the Indian stock market based on this outcome. Additionally, asset division can be utilized to signal systemic financial problems. Kocheturov et al. (2014), for instance, built a correlation matrix of stock returns for the US and Sweden, discovering that the stability of clusters might foretell the onset of a financial crisis. Therefore, it is vital to rationally divide the forked-coin market based on the daily return series.

The research above provides a great reference and maintains our interest in blockchain fork. However, few studies have examined the market structure for forked coins. Exploring the correlation relationship between forked coins serves three purposes: to help investors realize a diversified allocation of assets, provide macro-level structural information to the supervisor, and promptly identify the potential risk of the forked coins' market.

3. Data and methodology

3.1 Data

Although several BTC forks exist, only a restricted number of forked coins may collect information on closing prices.[2] Therefore, we used the fork list provided by the website, forkdrop.io, as the benchmark and finally obtained the daily closing price of BTC and its 14 forked coins from the website, coinmarketcap.com. Technical information about forked coins was obtained mainly from websites such as forkdrop.io (https://forkdrop.io/).

After determining the list of forked coins to be investigated and collecting the related closing price data, we must perform some statistical preprocessing of these sequences to scientifically interpret our results.

Price series in the financial sector are typically nonstationary (Sewell, 2011; Chen and Spokoiny, 2015), inducing pseudo-regression. To resolve this issue, we employ the logarithmic return according to Campbell et al. (2012)'s methodology. The formula for calculating the logarithmic rate of return for the 14 forked coins is as follows:

$$r_{i,t} = \ln\left(p_{i,t}/p_{i,t-1}\right), \tag{1}$$

where $p_{i,t}$ represents the closing price of the i-th fork on the t-th day; the closing price

[2] Regarding the relevant information on each forked coin, the authors collected and sorted it from various websites.

is in USD.

3.2 Methodology

3.2.1 Filter out the impact of BTC

In recent years, the clustering algorithm has become an essential tool for assessing market structure (Musmeci et al., 2015; León et al., 2017; Song et al., 2019). For example, using hierarchical clustering and the minimum-spanning-tree method, Song et al. (2019) analyzed the structure of the cryptocurrency market. They described the changes in the structure of the cryptocurrency market across multiple periods of regulatory intensity. Based hereon, we first eliminate BTC's influence to examine the market structure for the forked coins. Because BTC forked coins are forked from BTC, they are highly subject to price fluctuations in BTC. If the same factor influences these cryptocurrencies, directly grouping them may produce erroneous results (Song et al., 2019).

First, we use the maximal information coefficient (MIC) index (Reshef et al., 2011) and Pearson correlation coefficient to determine whether BTC and forked coins are correlated. A nonparametric approach derives MIC, and its value may show the correlation between variables. The Pearson correlation coefficient can indicate the linear correlation between variables, and researchers can employ statistical tests to test its significance. The following formula calculates the MIC between BTC and its forked coins *(Forks)*:

$$MIC[BTC; Forks] = \max_{|BTC||Forks|<B} \frac{I[BTC;Forks]}{log_2(\min(|BTC|,|Forks|))} \tag{2}$$

$$I[BTC; Forks] = \sum_{BTC,Forks} P(BTC, Forks)\, log_2 \frac{P(BTC,Forks)}{P(BTC)P(Forks)} \tag{3}$$

where *I[BTC;Forks]* represents the mutual information between the BTC and its forked coins. *P(BTC, Forks)* represents the joint probability density function of BTC and its forked coins, and *P(BTC)* and *P(Forks)* represent the marginal probability density functions of BTC and its forked coins. The value of parameter B is generally set to 0.6. The value of MIC ranges from 0 to 1, and the stronger the correlation between variables, the larger the corresponding value. MIC is a nonparametric statistic; thus, we cannot test its significance.

The Pearson correlation coefficient, ρ_i, is used to represent the linear correlation between the logarithmic return of the forked coin, R_i, and the BTC logarithmic return rate, $R_{bitcoin}$. The corresponding representation is as follows:

$$\rho_i = \frac{T\Sigma r_{it} r_{bitcoin,t} - \Sigma r_{it} \Sigma r_{bitcoin,t}}{\sqrt{T\Sigma r_{it}^2 - (\Sigma r_{it})^2}\sqrt{T\Sigma r_{bitcoin,t}^2 - (\Sigma r_{bitcoin,t})^2}} \quad t = 1,2,\dots,T; i \neq j \qquad (4)$$

where r_{it} represents the sample observation value of the logarithmic return rate of the forked coin, *i,* at time *t*. Next, we test the significance of the Pearson correlation coefficient. In addition, T represents the sample size.

$H_0\ \rho_i = 0$; $H_1\ \rho_i \neq 0$

The test statistic is given by:

$$t_i = \frac{\rho_i}{\sqrt{\frac{1-\rho_i^2}{T_i - 2}}} \qquad (5)$$

The corresponding degree of freedom is given by:

$$v = \mathrm{T} - 2 \qquad (6)$$

Several studies have examined BTC's substantial impact on its forked coins, focusing on lag and nonlinear situations. For instance, Yi et al. (2021) conducted a Granger causality test, including lagging periods 1 and 2. They showed that BTC lag had a statistically significant Granger causation to BCH. Additionally, they indicated in the study outlook that transfer entropy and nonlinear connection might be employed to analyze the information-flow transfer of forked coins. Therefore, we use the following model to eliminate BTC's impact on the forked coins:

$$r_{it} = \widehat{\alpha_i} + \widehat{\beta_{i0}} r_{bitcoin,t} + \widehat{\beta_{i1}} r_{bitcoin,t-1} + \widehat{\beta_{i2}} r_{bitcoin,t-2} + \widehat{\beta_{i3}} r_{bitcoin,t-3} + \widehat{\beta_{i4}} r_{bitcoin,t}^2 + \widehat{\beta_{i5}} r_{bitcoin,t}^3 + e_{it} \quad (7)$$

In Equation 7, r_{it} is the logarithmic return of BTC fork i at time t; $r_{bitcoin,t}$ $r_{bitcoin,t-1}$, $r_{bitcoin,t-2}$, and $r_{bitcoin,t-3}$ represent the sample observations of the logarithmic return on the same day, lag one day, lag two days, and lag three days, respectively. $r_{bitcoin,t}^2$ and $r_{bitcoin,t}^3$ represent the quadratic and tertiary terms of BTC's daily logarithmic return, respectively, representing the nonlinear impact of BTC on the fork. We select the variables according to the Bayesian information criterion (BIC) and modify the model (Equation (7)) based on the results of the lag selection. Thus, the derived regression residuals represent the returns without the effect of BTC.

Among them, $\left(\widehat{\alpha_i}, \widehat{\beta_{i0}}, \widehat{\beta_{i1}}, \widehat{\beta_{i2}}, \widehat{\beta_{i3}}, \widehat{\beta_{i4}}, \widehat{\beta_{i5}}\right)^T$ is the estimated value of intercept term and coefficients in the model (Eq. [7]). $\boldsymbol{e_{i,t}}$ is the residual of the model (Eq. (7)), which represents the return of the BTC fork after filtering the influence of BTC. r_{it} represents the sample observation value of the daily logarithmic return of the BTC fork. $r_{bitcoin,t-1}$, $r_{bitcoin,t-2}$, and $r_{bitcoin,t-3}$ represent the sample observations of the logarithmic return on the same day, lag one day, lag two days, and lag three days, respectively; while $r_{bitcoin,t}^2$ and $r_{bitcoin,t}^3$ represent the quadratic and cubic terms of BTC logarithmic return on that day, respectively.

3.2.2 Distance measurement of time series based on the PIC method

Measuring the distance between different time series is the premise and foundation of time series clustering. This study uses the method proposed by Piccolo (1990), which produces consistent estimators. Because the calculation is reasonably straightforward, it is often employed. Assuming there are two-time series $\boldsymbol{X}_T$ and $\boldsymbol{Y}_T$, the exact computation procedure is as follows:

$$\boldsymbol{X}_T = (X_1, \dots, X_T)^T$$

$$\boldsymbol{Y}_T = (Y_1, \dots, Y_T)^T$$

$\widehat{\Pi}_{X_T} = \left(\hat{\pi}_{1,X_T}, \dots, \hat{\pi}_{k_1,X_T}\right)^T$ and $\widehat{\Pi}_{Y_T} = \left(\hat{\pi}_{1,Y_T}, \dots, \hat{\pi}_{k_2,Y_T}\right)^T$ represent the value of the parameter in $AR(k_1)$ and $AR(k_2)$. The following formula calculates the distance between the two series.

$$d_{PIC}(X_T, Y_T) = \sqrt{\sum_{j=1}^{k}\left(\hat{\pi}'_{j,X_T} - \hat{\pi}'_{j,Y_T}\right)^2} \tag{8}$$

where Ω is a matrix of the weight.

$$k = \max(k_1, k_2) \tag{9}$$

$$\hat{\pi}'_{j,X_T} = \begin{cases} \hat{\pi}_{j,X_T}, j \le k_1 \\ 0, otherwise \end{cases} \tag{10}$$

$$\hat{\pi}'_{j,Y_T} = \begin{cases} \hat{\pi}_{j,Y_T}, j \le k_2 \\ 0, otherwise \end{cases} \tag{11}$$

3.2.3 Bottom-up hierarchical clustering

After determining the distance between multiple time series, the forked coins are clustered based on the dissimilarity matrix. Hierarchical clustering, K-means clustering, relocation clustering, and SOMs are the primary clustering techniques (Liao, 2005). Given that hierarchical clustering does not require the starting class center point to be specified and is insensitive to the sample input order, hierarchical clustering is used to examine fork market structure.

In our research, we employ the Ward method, developed by Ward (1963). Unlike other hierarchical clustering algorithms, the Ward method does not alter the members of the established clusters. Explaining the loss in the clusters is straightforward because the algorithm is not a black box. We combine classes using the Ward minimum variance method, which minimizes the increase in the sum of squared deviations within each class. We conduct the cluster analysis using the R statistical software package "Tsclust"

(Montero and Vilar, 2015) using the "ward.D2" approach specified by Murtagh and Legendre (2014). First, *n* samples are in a class, and each class merging increases the sum of squared deviations. The two merged categories of the fork that minimize the increase in the sum of squared deviations are chosen.

4. Results and discussion

4.1 Filter out the impact of BTC on forked coins

Considering that BTC may influence the clustering outcomes for its forked coins, we first evaluate the Pearson correlation coefficient and MIC between BTC and forked-coin return series to determine linear and nonlinear correlations. Based hereon, we attempt to use polynomials to characterize the relationship between BTC and forked coins and then analyze the residual series.

4.1.1 Correlations between BTC and its forked coins

To determine whether it is essential to reduce the impact of BTC in the series of forked coins based on data, we must first determine whether a correlation exists between BTC and the forked coins, including linear and nonlinear correlation. Therefore, we attempt to describe the correlation between series using the MIC index. We assess the significance of the Pearson correlation coefficient to establish whether there is a substantial linear correlation between BTC and its forked coins.

The MIC and Pearson values represent the correlation between BTC and its forked coins (see Table 1). Among these, the MIC value shows the entire correlation of returns, including linear and nonlinear relationships. MIC is the correlation coefficient derived in a nonparametric form, and its value ranges from 0 to 1. According to Lin et al. (2012), a correlation between variables exists when the MIC value is more than 0.02. Table 1 shows that BCH and Bitcoin Gold (BTG) have the highest connection with BTC, but

United Bitcoin (UBTC) and Bitcoin File (BIFI) have lower MIC values. In addition, the Pearson correlation coefficient data indicate significant linear relationships between BTC and its forked coins. Consequently, the impact of BTC on the return on its forked coins has become a crucial aspect that must be considered. We calculate MIC and Pearson correlation coefficients to demonstrate that the return on forked coins is closely tied to that on BTC, providing quantitative justification for filtering BTC's influence.

<Insert Table 1 Here>

4.1.2 Filtering out BTC's influence on forked coins

Table 1 demonstrates the influence of BTC on its forked coins. To avoid pseudo-regression, we conduct a unit-root test on the logarithmic return series before filtering out BTC's influence. The results of the unit-root inspection are shown in Table 2. The ADF test demonstrates that each return series is stationary; therefore, we may utilize Equation (7) to filter out BTC's influence.

<Insert Table 2 Here>

For variable selection, we employ the BIC criterion to prevent variable redundancy. Figure 2 depicts the selection of variables based on the BIC criterion, which comprises 14 fundamental images correlating to 14 forked coins. The horizontal axis shows the seven variables involved in variable selection, while the vertical axis indicates the BIC value under various variable selection criteria. In variable selection, the optimal variable combination has the smallest BIC value. In Figure 2, this variable corresponds to the darkest possible color combination.

<Insert Figure 2 Here>

It is worth emphasizing that BTC's nonlinear and time-delay effects cannot be ignored, for example, the nonlinear influence of BTC on BCH, BTG, and Bitcoin SV (BSV) return and the time-delay effect of BTC on Bitcoin God (GOD) and Bitcoin 2

(BTC2). We employ the model residual as research data to be evaluated next; Table 3 presents the determined model fitting according to the BIC criteria.

<Insert Table 3 Here>

Figure 3 reflects the correlation between BTC and its forked coins before and after filtering out BTC. The Pearson correlation values determine the grid's color; red represents a negative correlation, whereas blue represents a positive correlation. The darker the color, the greater the absolute number. The values within the grid correspond to the Pearson correlation coefficients. Furthermore, we provided comparable scores to grids that fail the significance test. Positive correlations exist between forked coins, and the correlation relationship decreases when the BTC influence is filtered out. Even after excluding BTC, a considerable positive correlation remains between BCH, BCD, BTG, and BSV.

<Insert Figure 3 Here>

4.2 Clustering results after filtering out BTC's influence

4.2.1 Basic clustering results

Before hierarchical clustering, it must be determined whether the samples should be divided into many clusters. Thus, the elbow rule is used to determine the number of clusters that have been subdivided. Figure 4 illustrates that the horizontal axis represents the number of clusters into which the sample has been divided, while the ordinate axis represents the sum of the squares of the deviations inside the cluster.

<Insert Figure 4 Here>

This study separates the forked coins into three clusters following the elbow rule, number of clusters, and notion of minimizing intra-group error. Figure 5 illustrates the results of the hierarchical clustering of forked coins within and outside BTC. The clustering results are consistent and independent of the BTC filtering.

<Insert Figure 5 Here>

Cluster 1 comprises five forked coins: GOD, BCX, Super Bitcoin (SBTC), BIFI, and UBTC. On January 11, 2018, GOD forked from BTC and was listed on three cryptocurrency exchanges. GOD will airdrop 17 million tokens to current BTC accounts based on the number of BTCs mined at the time of listing; the remaining 4 million tokens will be mined by proof-of-stake (PoS) miners , which will be utilized for charity airdrops. BCX forked at a block height of 4,498,888. The BCX project introduced smart contracts and proofs requiring no prior knowledge and the functions of unique addresses and replay protection. In addition, BCX utilizes delegated proof of stake (dPoS) to increase accounting efficiency, decrease transaction costs, and conserve energy. SBTC was issued for the first time on December 5, 2017; the project is now listed on two exchanges and was forked at a block height of 498,888. SBTC will improve scalability, privacy protection, and block capacity compared to BTC, and the initial BTC account will receive the same quantity of airdrops.[3] BIFI was issued for the first time on December 27, 2017. The project adopted BTC's settlement network with the InterPlanetary File System (IPFS) interplanetary file system to construct a distributed-network file system and resolve the mining process' energy consumption issue. The fork time for UBTC is unknown; UBTC is a forked coin in the form of passive or registered airdrops dedicated to creating BTC software and safeguarding users against fraud.

Cluster 2 comprises BTC and four forked coins: BCD, BTG, BCH, and BSV. The first instance of BCD issuance occurred on November 16, 2017, at a block height of 495,866. The project is currently listed and traded on 17 exchanges. The project keeps

---

[3] Data source：https://www.coinsuggest.com/super-bitcoin-sbtc/

the BTC blockchain data before the fork and creates the blockchain using a new work-verification algorithm. BCD increases the block capacity to 8Mb and sets the block size to dynamic mode; the total quantity of BCD is ten times that of BTC, which reduces the user's participation threshold and effort to some level. In addition, the BCD blockchain uses Segregated Witness and Lightning Networks. The date of BTG's initial issuance was October 24, 2017. It forked at a BTC block height of 491,407, and 15 exchanges listed and traded the project. BTG, like BCH, implements a pre-mining process. The project implements a novel PoW algorithm, complete replay protection, and a unique wallet address. The project is mainly utilized for value preservation under inflation and cross-border remittances. The BCH's first release date was July 24, 2017. BCH is presently listed and traded on 115 exchanges, and it has expanded its on-chain capacity relative to the BTC blockchain. The maximum block size is increased to 8Mb, and a new technique for hash signature and difficulty adjustment is introduced. The project is primarily utilized for small business transactions and payment processing; BSV was issued for the first time on November 9, 2018, and forked from BCH. BSV has developed a novel full-node BCH implementation technique. BSV aims to achieve extensive on-chain expansion.

Cluster 3 includes five forked coins: MicroBitcoin (MBC), Bitcoin Atom (BCA), Bitcoin Interest (BCI), CLAMs（CLAM）, and Bitcore (BTX). On October 27, 2018, MBC became publicly traded. It aims to become a flexible means of payment for clients. The issuance of MBC is 10,000 times the total BTC supply, and LWMA-3[4] is used as the new encryption algorithm. The first BCA was issued on January 12, 2018. BCA was forked during a peak of BTC of 505,888. The project's total value is 21 million pieces,

[4] https://github.com/zawy12/difficulty-algorithms/issues/3

currently listed on an exchange. It supports the Lightning Network, Segregated Witness, and atomic swaps across chains. BCA employs a hybrid consensus mechanism, which decreases the likelihood of attacks by 51%. The first BCI was issued on May 3, 2018. Maximum availability is 22.3 million. It is an investment-based cryptocurrency focused primarily on technology, the community, and savings. It is a decentralized platform for paying interest. The first BTX was issued on April 24, 2017. Currently, it is listed and traded on three exchanges. It employs Segregated Witnesses and enhances the blockchain's capacity. The total issuance is 21 million, and the initiative aims to become an efficient and low-cost payment instrument. BTC2 is produced at a block height of block 507,850, and the original BTC account can receive the same number of tokens. The project did not directly replicate the BTC source code but utilized the snapshot-fork method. In addition, the project uses zero-knowledge proof, aiming to create cash-like digital cash that is fast, safe, decentralized, and private. CLAM appears to be a passive airdrop committed to fixing the life-cycle problem of BTC, consequently expediting the main chain's extension process.

4.2.2 Comparison of three clusters

After removing BTC's influence, we compare the characteristics of the logarithmic return series and forking backdrop. Table 4 summarizes the statistical characteristics of the three forked-coin clusters. The average returns on the three clusters are all negative from the standpoint of the average level. Cluster 1 represents the most severe loss, followed by cluster 3 and then cluster 2. Furthermore, based on the characteristics of the distributions of logarithmic returns, all three clusters exhibit a right deviation and a peak with a thick tail. Cluster 2 is the most evident performer among them.

<Insert Table 4 Here>

Moreover, comparing the distribution of extreme values reveals that the difference between the largest and the least values in cluster 2 is the smallest. Finally, the fluctuation characteristics of the three clusters are computed. After analyzing the standard deviation, we determine that cluster 1 has the biggest fluctuation range, followed by cluster 3 and then cluster 2, which has the smallest fluctuation. After analyzing the returns in three clusters, cluster 1 has the lowest average return. Cluster 2 has the lowest loss and fewest fluctuations. Cluster 3's performance falls between those of clusters 1 and 2.

We next investigate the technical differences between the three clusters, as shown in Table 5. The first is the forking time and height of its blocks. The median forking times for the three clusters are 2017-12-19, 2017-9-12, and 2018-1-24, with corresponding median forking heights of 500,056, 5,484,982, and 5,505,888. Cluster 2 is the oldest, followed by cluster 1 and then cluster 3.

<Insert Table 5 Here>

Regarding the consensus technique, PoW is the most often utilized process in forks, which BTC also employs. In cluster 1, SBTC, UBTC, and BCX all utilize PoW, with SBTC and UBTC utilizing SHA256 and BCX utilizing Blake2. GOD and BIFI both utilize the PoS consensus. Then, each forked coin in cluster 2 adopts PoW. BCD is based on the X13 algorithm, BTG on the Equihash method, and BCH and BSV on SHA256. In cluster 3, BTC2 and CLAM use PoS; MBC and BTX use PoW based on Grostl and Timetravel10, and BCA uses hybrid consensus. Moreover, in terms of the codebase, most forks use the BTC Core code, whereas BCH uses the BTC Clashic code and BSV uses the BCH ABC code. Regarding the codebase, most forks adopt the BTC Core code; however, in cluster 2, BCH uses BTC Clashic, and BSV uses BCH ABC coding. In addition, we investigate whether these forks support SegWit, one of the ideas for executing a BTC soft fork and bypassing the blockchain size restriction. Bech32, one of the SegWit addresses, is utilized by BCD, BTG, and BCA.

Overall, by analyzing the technical parameters of each cluster, we find that forks in cluster 2 maintained the PoW employed by BTC but with slight modifications to the algorithm. In addition to SHA256, cluster 2 forks introduce the X13 and Equihash algorithms. The code bases for BCH and BSV, which still employ the SHA256 algorithm, are altered to BTC classic and BCH ABC. Therefore, forks in cluster 2 are more technologically innovative; however, all the forks in cluster 1 are based on the BTC Core code base, and no code innovation has occurred. In the consensus procedure, cluster 1 forks are different. GOD and BIFI employ PoS, whereas others employ PoW. Cluster 1 has a later forking time than cluster 2, although both clusters support SegWit. Therefore, we refer to cluster 1 as the forked coin that supports SegWit. Cluster 3 has the latest forking time, lowest median price, and fewest futures among the three clusters. Cluster 3 forks all utilize BTC Core, although only BCA implements SegWit with address bech32. There are substantial disparities between the three clusters in terms of return performance and technical solutions; thus, we refer to cluster 3 as the latest forked coin. As the early forks, cluster 2 has a high level of technical innovation, the highest price, and the most futures; therefore, we refer to it as the mature forked coin. Cluster 3 is the latest to fork, has the lowest price and return, least innovative underlying code, and has more diverse consensus processes.

4.3 Further analysis

In order to analyze the changes of clusters formed by BTC and forked coins over time, we employ cluster analysis on samples of forked coins by year. Figure 6 displays the clustering results, showing that several forked coins did not constitute a different cluster in the early days. Nonetheless, by 2019, three clusters had formed, with the slow expansion of forks and rapid development of early forks. The results in 2020 are similar to the outcomes of clustering the complete sample.

<Insert Figure 6 Here>

It is important to note that the clustering results are not immutable. SBTC and BCX, for instance, belonged to cluster consisting of BCA MBC BTX BCX SBTC in 2019 and to the cluster consisting of GOD BIFI BCX SBTC in 2020.Transferring forked coins between clusters demonstrates that the market structure for forked coins is unstable. We also find that by 2021, the clustering of BTC and forked coins will shift from three cluters to two clusters. In addition, we discover that the well-performing forked coins remained in the same cluster as BTC over time.

5. Conclusions and policy recommendations

With the introduction of blockchain, forked coins have created a new financial market that is an essential addition to the traditional and cryptocurrency markets. Thus, the forked- coin market has recently been a significant concern. In this instance, investors can choose to invest in forked coins, and BTC holders can receive a payout in forked coins proportional to their holdings. However, a digital asset carries risks that cannot be overlooked, as they may lead to risk accumulation and possibly systemic risk.

Our study is the first to conduct an economic and technical analysis of the forked-coin market. From 2018 through 2021, this study examine BTC's market-structure characteristics, changes, and forked coins. After removing the linear, nonlinear, and time-lag effects from BTC's influence on the forked coins, we construct a polynomial regression model and derived the residuals. The residuals are then used as the study sequence, hierarchical clustering was performed from the bottom up, and the coins were separated into three clusters. Our research findings are the following: (1) BTC has a considerable effect on the return on the forked coins but none on the clustering result. Therefore, BTC is not a determining element in the market structure for the forked coins, whereas the characteristics of the forked coins are. (2) Recently, forked coins have manifested into three clusters: SegWit-supported forked coins, mature forked coins, and the latest forked coins. The earliest forked coins from the BTC chain perform the best. Their financial offerings are the most advantageous. (3) Their financial products are relatively single for the latest forked coins, and their return are significantly less than that of BTC and prior forked coins.These coins with a fork can be classified into three clusters, each of which has its own characteristics. (4) The market structure of BTC and forked coins is not immutable. On the one hand, individual forked coins (such as BCX SBTC) will belong to different clusters in different years; On the other hand, the appropriate number of clusters will also take different values at different years.

Our findings offer investors and governments new insights into the investing strategy and market for forked coins and offer investors a reference for their investment strategies. First, in the three clusters, the forked coins in the same cluster as BTC (such as BCH, BTG, BSV, and BCD) have more financial products and a higher closing price than other forked coins; thus, their performance is superior to that of other forked coins. Second, because the performance of the three clusters differs vastly, investment in forked coins must also adhere to the maxim, “Don’t put all your eggs in one basket.” Therefore, investors should focus on different clusters. Finally, investors should also pay attention to the change of the market structure over time. For those forked coins that leave the BTC cluster over time, investors should do a good job in risk control.Our research assists policymakers in understanding the market features of forked coins from a macro perspective. Based on the conclusion, financial regulatory authorities can implement risk-prevention measures for forked coins based on their unique characteristics. Cluster 1 consists of forked coins with low returns and many derivatives. Credit risk should receive special attention from the regulatory authorities. Due to the absence of reasonably consistent project objectives, regulator must pay special attention to the market risk associated with passive airdrop projects exhibiting similar return characteristics. Regulators should also pay attention to the forked coins that are transferred between different clusters over time, especially those that are removed from the BTC cluster.

Declarations

Conflicts of interest

The authors have no relevant financial or nonfinancial interests to disclose.

Reference

Al Mamun, M., Uddin, G. S., Suleman, M. T., Kang, S. H., 2020. Geopolitical risk, uncertainty and Bitcoin investment. *Phys. A: Stat. Mech. Appl.* 540, 123107. https://doi.org/10.1016/j.physa.2019.123107.

Aste, T., 2019. Cryptocurrency market structure: connecting emotions and economics. *Digit. Finance* 1(1), 5-21. https://doi.org/10.1007/s42521-019-00008-9.

Bazán-Palomino, W., 2020. Bitcoin and its offspring: a volatility risk approach. Springer, Singapore. https://doi.org/10.1007/978-981-15-4498-9_13.

Bazán-Palomino, W., 2021. How are Bitcoin forks related to Bitcoin? *Finance Res. Lett.* 40, 101723. https://doi.org/10.1016/j.frl.2020.101723.

Biais, B., Bisiere, C., Bouvard, M., Casamatta, C., 2019. The Blockchain folk theorem. *Rev. Financ. Stud.* 32(5), 1662-1715. https://doi.org/10.1093/rfs/hhy095.

Bouri, E., Shahzad, S. J. H., Roubaud, D., Kristoufek, L., Lucey, B., 2020. Bitcoin, gold, and commodities as safe havens for stocks: new insight through wavelet analysis. *Q. Rev. Econ. Finance* 77, 156-164. https://doi.org/10.1016/j.qref.2020.03.004.

Bowden, C., 2018. Forking in time. *A Peer Rev. J. About* 7(1), 140-150. https://doi.org/10.7146/aprja.v7i1.116061.

Cagli, E. C., 2019. Explosive behavior in the prices of Bitcoin and altcoins. *Finance Res. Lett.* 29, 398-403. https://doi.org/10.1016/j.frl.2018.09.007.

Campbell, J. Y., Lo, A. W., MacKinlay, A. C., 2012. The econometrics of financial markets. Princeton University Press, Princeton. https://doi.org/10.1515/9781400830213.

Chaim, P., Laurini, M. P., 2018. Volatility and return jumps in bitcoin. *Econ. Lett.* 173, 158-163. https://doi.org/10.1016/j.econlet.2018.10.011.

Chason, E. D. ,2019a. A Tax on the Clones: The Strange Case of Bitcoin Cash. *Va. Tax Rev.* 39, 1. Available at: https://scholarship.law.wm.edu/facpubs/1952.

Chason, E. D., 2019b. Cryptocurrency Hardhard forks and revenue ruling 2019-24. *Va. Tax Rev.*,. 39, 279. Available at: https://scholarship.law.wm.edu/facpubs/1995.

Chen, Y., Spokoiny, V. 2015. Modeling nonstationary and leptokurtic financial time series. *Econ Theory* 31(4), 703-728. https://doi.org/10.1017/S0266466614000528.

Ciaian, P., Rajcaniova, M., 2018. Virtual relationships: short-and long-run evidence from BitCoin and altcoin markets. *J. Int. Financial Mark. Inst. Money* 52, 173-195. https://doi.org/10.1016/j.intfin.2017.11.001.

Cotler, B., 2020. Cryptocurrency tax update: if there's a hard fork in the road, take it (or not). *J. Tax. Investm.* 37(2), 43-48.

Courtois, N. T., Bahack, L., 2014. On subversive miner strategies and block withholding attack in bitcoin digital currency. arXiv preprint arXiv:1402.1718. http://arxiv.org/abs/1402.1718.

Demir, E., Simonyan, S., García-Gómez, C. D., Lau, C. K. M., 2021. The asymmetric effect of bitcoin on altcoins: evidence from the nonlinear autoregressive distributed lag (NARDL) model. *Finance Res. Lett.* 40, 101754. https://doi.org/10.1016/j.frl.2020.101754.

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

Fauzi, M. A., Paiman, N., Othman, Z., 2020. Bitcoin and cryptocurrency: challenges, opportunities and future works. *J. Asian Finance Econ. Bus.* 7(8), 695-704. https://doi.org/10.13106/JAFEB.2020.VOL7.NO8.695.

Göbel, J., Krzesinski, A. E., 2017. Increased block size and Bitcoin blockchain dynamics. *27th International Telecommunication Networks and Applications Conference (ITNAC)*, 1:6. https://doi.org/10.1109/ATNAC.2017.8215367.

Hsu, T. H., Chu, K. M., Chan, H. C. 2000. The fuzzy clustering on market segment. *IEEE Fuzzy Syst.* 2, 621-626. https://doi.org/10.1109/FUZZY.2000.839064.

Hu, J., Härdle, W. K., Kuo, W., 2019. Risk of Bitcoin market: volatility, jumps, and forecasts. arXiv preprint arXiv:1912.05228. https://doi.org/10.48550/arXiv.1912.05228.

Iansiti, M., Lakhani, K., 2017. The truth about blockchain. *Harv. Bus. Rev.* 1-11. Available at: https://e-tarjome.com/storage/btn_uploaded/2019-09-25/1569393941_10128-etarjome-English.pdf.

Islam, A., Mantymaki, M., Turunen, M., 2019. Why do blockchains split? An actor-network perspective on Bitcoin splits. *Technol. Forecast Soc. Change* 148(11), 119743. https://doi.org/10.1016/j.techfore.2019.119743.

Johnson, J., 2021. Does Bitcoin cash have a mind of its own, or will Bitcoin always determine its future?. Available at SSRN: https://ssrn.com/abstract=3818077 or http://dx.doi.org/10.2139/ssrn.3818077.

Kocheturov, A., Batsyn, M., Pardalos, P. M., 2014. Dynamics of cluster structures in a financial market network. *Phys. A: Stat. Mech. Appl.* 413, 523-533. https://doi.org/10.1016/j.physa.2014.06.077.

Kubiak, P., Kutylowski, M., 2020. Preventing a fork in a Blockchain - David fighting Goliath. *IEEE Security Privacy* https://doi.org/10.1109/TrustCom50675.2020.00139.

Kurka, J., 2019. Do cryptocurrencies and traditional asset classes influence each other? *Finance Res. Lett.* 31, 38-46. https://doi.org/10.1016/j.frl.2019.04.018.

León, C., Kim, G.-Y., Martínez, C., Lee, D., 2017. Equity markets' clustering and the global financial crisis. *Quant. Financ.* 17(12), 1905-1922. https://doi.org/10.1080/14697688.2017.1357970.

Li, Z. Z., Tao, R., Su, C. W., Lobonţ, O. R., 2019. Does Bitcoin bubble burst? *Qual. Quant.* 53(1), 91-105. https://doi.org/10.1007/s11135-018-0728-3.

Liao, T. W., 2005. Clustering of time series data—a survey. *Pattern Recognit.* 38(11), 1857-1874. https://doi.org/10.1016/j.patcog.2005.01.025.

Lin, C., Miller, T., Dligach, D., Plenge, R., Karlson, E., Savova, G., 2012, July. Maximal information coefficient for feature selection for clinical document classification. *In ICML workshop on machine learning for clinical data. Edinburgh, UK.* Available at: https://people.cs.pitt.edu/~milos/icml_clinicaldata_2012/Papers/Poster_Chen_Guergana_ICML_Clinical_2012.pdf.

Lin, I. C., Liao, T. C., 2017. A survey of blockchain security issues and challenges. *Int.*

J. Netw. Secur. 19(5), 653-659. https://doi.org/10.6633/IJNS.201709.19(5).01.

Liu, Q. L., Xu, Y. Y., Cao, B., Zhang, L., Peng, M. G., 2021. Unintentional forking analysis in wireless blockchain networks. *Digit. Commun. Netw.* 7(3), 335-341. https://doi.org/10.1016/j.dcan.2020.12.005.

Lu, X., Liu, C., Lai, K. K., Cui, H., 2021. Risk measurement in Bitcoin market by fusing LSTM with the joint-regression-combined forecasting model. *Kybernetes*, (ahead-of-print). https://doi.org/10.1108/K-07-2021-0620.

Misic, V. B., Misic, J., Chang, X. L., 2019. On forks and fork characteristics in a Bitcoin-like distribution network. *IEEE Blockchain.* https://doi.org/10.1109/Blockchain.2019.00035.

Momeni, M., Mohseni, M., Soofi, M., 2015. Clustering stock market companies via k-means algorithm. *Arab. J. Bus. Manag. Rev.* 33(2578), 1-10. https://platform.almanhal.com/Files/Articles/75353.

Montero, P., Vilar, J.A., 2015. TSclust: an R package for time series clustering. *J. Stat. Softw.* 62(1), 1-43. Available at: http://www.jstatsoft.org/v62/i01/.

Möser, M., Soska, K., Heilman, E., Lee, K., Heffan, H., Srivastava, S., Hogan, K., Hennessey, J., Miller, A., Narayanan, A., 2017. An empirical analysis of traceability in the monero blockchain. *arXiv preprint arXiv:1704.04299*. https://doi.org/10.48550/arXiv.1704.04299.

Murtagh, F., Legendre, P., 2014. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *J. Classif.* 31(3), 274-295. https://doi.org/10.1007/s00357-014-9161-z.

Musmeci, N., Aste, T., & Di Matteo, T., 2015. Relation between financial market structure and the real economy: comparison between clustering methods. *PloS One*, 10(3), e0116201. https://doi.org/10.1371/journal.pone.0116201.

Nakamoto, S., 2008. Bitcoin: a peer-to-peer electronic cash system. *Decentralised Bus. Rev.* 21260. Available at: https://assets.pubpub.org/d8wct41f/31611263538139.pdf.

Nanda, S. R., Mahanty, B., Tiwari, M. K., 2010. Clustering Indian stock market data for portfolio management. *Expert Syst. Appl.* 37(12), 8793-8798. https://doi.org/10.1016/j.eswa.2010.06.026.

Nyman, L., Mikkonen, T., Lindman, J., Fougere, M., 2012. Perspective of code forking and sustainability in open source software. *8th IFIP WG 2.3 International Conference* 378, 274-279. https://doi.org/10.1007/978-3-642-33442-9_21ff. ffhal-01519038f.

Piccolo, D., 1990. A distance measure for classifying ARIMA models. *J. Time Ser. Anal.* 11(2), 153-164. https://doi.org/10.1111/j.1467-9892.1990.tb00048.x.

Poon, J., Dryja, T., 2016. The Bitcoin lightning network: scalable off-chain instant payments. Available at: https://lightning. network/lightningnetwork-paper. pdf.

Reshef, D. N., Reshef, Y. A., Finucane, H. K., Grossman, S. R., McVean, G., Turnbaugh, P. J., Lander, E. S., Mitzenmacher, M., Sabeti, P. C., 2011. Detecting novel associations in large data sets. *Science* 334 (6062), 1518-1524. https://doi.org/10.1126/science.1205438.

Schär, F., 2020. Blockchain forks: a formal classification framework and persistency analysis. *Singap. Econ. Rev.* 1-11. https://doi.org/10.1142/S0217590820470025.

Sewell M. , 2011. Characterization of financial time series. Research Note RN/11/01. UCL Department of Com-puter Science, London. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=ecb38344a3d16f255a7a7edc0ba27adebaa3d9d6.

Song, J. Y., Chang, W., Song, J. W., 2019. Cluster analysis on the structure of the cryptocurrency market via Bitcoin-Ethereum filtering. *Physica A* 527, 121339. https://doi.org/10.1016/j.physa.2019.121339.

Tomić, N., 2020. Measuring the effects of Bitcoin forks on selected cryptocurrencies using event study methodology. *Industrija* 48(2), 21-36.https://doi.org/10.5937/industrija48-26003.

Troster, V., Tiwari, A. K., Shahbaz, M., Macedo, D. N., 2019. Bitcoin returns and risk: a general GARCH and GAS analysis. *Finance Res. Lett.* 30, 187-193. https://doi.org/10.1016/j.frl.2018.09.014.

Trucíos, C., 2019. Forecasting Bitcoin risk measures: a robust approach. *Int. J. Forecast.* 35(3), 836-847. https://doi.org/10.1016/j.ijforecast.2019.01.003.

Van Alstyne, M., 2014. Why Bitcoin has value. *Commun. ACM* 57(5), 30-32. https://doi.org/10.1145/2594288.

Vujičić, D., Jagodić, D., Ranđić, S., 2018. Blockchain technology, bitcoin, and Ethereum: A brief overview. 17th international symposium infoteh-jahorina (infoteh). https://doi.org/10.1109/INFOTEH.2018.8345547.

Ward Jr, J. H., 1963. Hierarchical grouping to optimise an objective function. *J. Am. Stat. Assoc.* 58(301), 236-244. https://doi.org/10.1080/01621459.1963.10500845.

Webb, N., 2018. A fork in the blockchain: income tax and the Bitcoin/Bitcoin cash hard fork. *N.C. J. Law Technol.* 19(4), 283. Available at: https://scholarship.law.unc.edu/ncjolt/vol19/iss4/10.

Xu, D., 2019. Free money, but not tax-free: A proposal for the tax treatment of cryptocurrency hard forks. *Actual Probs. Econ. & L.*, 1661. Available at: https://ir.lawnet.fordham.edu/flr/vol87/iss6/14.

Yi, E., Cho, Y., Sohn, S., Ahn, K., 2021. After the splits: information flow between Bitcoin and Bitcoin family. *Chaos Solit. Fractals* 142, 110464.

https://doi.org/10.1016/j.chaos.2020.110464.

Tables

Table 1. The correlation relationship between BTC and its forked coins

Forked coins	MIC	Pearson	t-value	P-value
BCH	0.4882	0.7751	46.835	0
BCA	0.1948	0.1979	7.6763	0
BCD	0.3264	0.4233	17.84	0
BIFI	0.1512	0.1401	5.0012	0
GOD	0.1936	0.0145	0.5516	0.5813
BTG	0.4336	0.6717	34.625	0
BSV	0.3424	0.5430	21.888	0
BCX	0.1627	0.1931	7.5153	0
BTC2	0.2225	0.2805	8.9889	0
BTX	0.1961	0.3387	13.743	0
CLAM	0.2522	0.1487	5.7415	0
UBTC	0.1324	0.0205	0.7833	0.4336
MBC	0.1578	0.1248	4.284	0
SBTC	0.1567	0.1643	6.3619	0

Note: The data used in calculating MIC and Pearson correlation coefficients are the logarithmic returns of BTC and its forked coins.

Table 2. Results of Augment Dickey–Fuller test on return series

Number	Forked coins	Dickey–Fuller	P-value	Stationary	Observation
1	BCH	−26.185	<0.01	yes	1461
2	BCA	−33.335	<0.01	yes	1448
3	BCD	−31.965	<0.01	yes	1461
4	BIFI	−37.773	<0.01	yes	1252
5	GOD	−37.692	<0.01	yes	1449
6	BTG	−27.012	<0.01	yes	1461
7	BSV	−25.565	<0.01	yes	1148
8	BCX	−39.726	<0.01	yes	1461
9	BTC2	−27.514	<0.01	yes	948
10	BTX	−31.698	<0.01	yes	1461
11	CLAMS	−36.669	<0.01	yes	1461
12	UBTC	−38.36	<0.01	yes	1461
13	MBC	−30.441	<0.01	yes	1162
14	SBTC	−41.222	<0.01	yes	1461
15	BTC	−26.657	<0.01	yes	1461

Note: When calculating Dickey–Fuller statistics, the lag order is 1.

Table 3. Analysis of BTC influence on the forked coins

Forked coins	BTC	BTC (−1)	BTC (−2)	BTC^2	BTC^3	Adjusted-R^2	F-test
BCH	1.2613*** [0.0291]			−1.6879*** [0.2983]	−3.7172*** [0.7904]	0.6084	0.0000
BCA	0.9079*** [0.1196]					0.0377	0.0000
BCD	1.1300*** [0.0631]					0.1798	0.0000
BIFI	0.9550*** [0.1912]					0.0188	0.0000
GOD		0.7531** [0.3635]				0.0023	0.0345
BTG	1.1545*** [0.0360]			−1.9796*** [0.3692]	−4.1429*** [0.9781]	0.4607	0.0000
BSV	1.1290*** [0.0574]			−2.1122*** [0.5759]	−3.9671*** [1.4864]	0.3015	0.0000
BCX	1.1654*** [0.1551]					0.0366	0.0000
BTC2	1.0161*** [0.1113]		−0.3417*** [0.1113]			0.0858	0.0000
BTX	0.9139*** [0.0666]					0.1137	0.0000
CLAM	0.8386*** [0.1462]					0.0214	0.0000
UBTC	0.7514*** [0.1115]					0.0295	0.0000
MBC	0.6142*** [0.1442]					0.0146	0.0000
SBTC	0.9780*** [0.1539]					0.0269	0.0000

Table 4. Quantitative characteristics among three different cluster

cluster	forks	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
	GOD	−0.0021	0.6651	−0.0035	−0.0050	0.0833	−3.2865	3.4496	6.7361	0.2072	6.4371	0.0216
	SBTC	−0.0046	0.2698	−0.0030	−0.0058	0.1617	−1.3134	1.4998	2.8132	0.1259	5.5565	0.0088
cluster1	BCX	−0.0037	0.2758	−0.0104	−0.0078	0.1014	−1.8629	1.6744	3.5374	0.4722	10.6452	0.0090
	UBTC	−0.0018	0.1618	−0.0076	−0.0064	0.0742	−1.1535	1.1057	2.2592	0.0766	11.9335	0.0053
	BIFI	−0.0035	0.2434	−0.0079	−0.0074	0.1099	−1.1821	1.3016	2.4838	0.1144	5.7843	0.0079
	BCD	−0.0017	0.0636	−0.0058	−0.0057	0.0339	−0.2576	0.6918	0.9494	3.0419	25.6327	0.0021
	BTG	0.0013	0.0513	−0.0036	−0.0029	0.0267	−0.2216	0.6386	0.8602	3.9258	39.1354	0.0017
cluster2	BCH	0.0002	0.0374	−0.0023	−0.0016	0.0239	−0.2228	0.3383	0.5612	1.9507	16.7881	0.0012
	BSV	0.0011	0.0573	−0.0045	−0.0029	0.0276	−0.2987	0.8106	1.1093	4.5930	55.4340	0.0019
	BTC	0.0018	0.0403	0.0014	0.0021	0.0263	−0.4647	0.1718	0.6366	−1.6090	19.9692	0.0013
	BTC2	−0.0007	0.1381	−0.0053	−0.0040	0.0749	−0.7420	1.9868	2.7287	3.2322	47.6355	0.0045
	CLAM	−0.0043	0.2741	−0.0019	−0.0036	0.0306	−2.1784	1.5168	3.6952	−0.0920	14.3347	0.0089
cluster3	MBC	−0.0027	0.1807	−0.0009	−0.0057	0.0962	−1.0512	1.2720	2.3232	0.4127	5.5697	0.0059
	BCA	−0.0024	0.1905	−0.0002	−0.0068	0.0985	−1.4431	1.0321	2.4752	0.0660	7.8594	0.0062
	BTX	−0.0035	0.1012	−0.0075	−0.0056	0.0601	−0.6190	0.6849	1.3039	0.4460	9.1340	0.0033
cluster1	mean	−0.0031	0.3232	−0.0065	−0.0065	0.1061	−1.7597	1.8062	3.5659	0.1993	8.0713	0.0105
cluster2	mean	0.0005	0.0500	−0.0029	−0.0022	0.0277	−0.2931	0.5302	0.8233	2.3805	31.3919	0.0016
cluster3	mean	−0.0027	0.1769	−0.0032	−0.0051	0.0721	−1.2067	1.2985	2.5052	0.8130	16.9067	0.0057

Table 5. Technical characteristics among three different clusters

Cluster	Forks	Consensus	Codebase	Bech32	SegWit	Height	Forked Time	Projected Type	Deposit	Futures
cluster1	GOD	PoS	Unknown	Unknown	yes	501225	2017/12/27	BTC Chain Fork	2	4
	SBTC	PoW SHA256	Bitcoin Core	Unknown	yes	498888	2017/12/12	BTC Chain Fork	5	13
	BCX	PoW Blake2	Bitcoin Core	Unknown	yes	498888	2017/12/12	BTC Chain Fork	5	9
	UBTC	PoW SHA256	Bitcoin Core	Unknown	yes	Unknown	Unknown	Passive Airdrop Registered Airdrop	7	7
	BIFI	PoS	Bitcoin Core	Unknown	yes	501225	2017/12/27	BTC Chain Fork	1	2
cluster2	BCD	PoW X13	Bitcoin Core	yes	yes	495866	2017/11/24	BTC Chain Fork	18	17
	BTG	PoW Equihash	Bitcoin Core	yes	yes	491407	2017/10/24	BTC Chain Fork	32	5
	BCH	PoW SHA256	Bitcoin Clashic	no	no	478558	2017/8/1	BTC Chain Fork	15	53
	BSV	PoW SHA256	Bitcoin Cash ABC	no	no	478558	2017/8/1	BTC Chain Fork	10	8
cluster3	BTC BTC2	PoS	Unknown	Unknown	Unknown	507850	2018/6/14	Passive Airdrop	2	0
	CLAM	PoS	Bitcoin Core	no	no	300377	2014/5/12	Passive Airdrop	6	0
	MBC	PoW Grostl	Bitcoin Core	Unknown	yes	525000	2018/5/28	BTC Chain Fork	5	0
	BCA	Hybrid PoS PoW SHA256	Bitcoin Core	yes	yes	505888	2018/1/24	BTC Chain Fork	4	1
	BTX	PoW Timetravel10	Bitcoin Core	Unknown	yes	492820	2017/11/2	Passive Airdrop	12	0

Note: The technical information of the forked currency is obtained according to the network data.

Figures

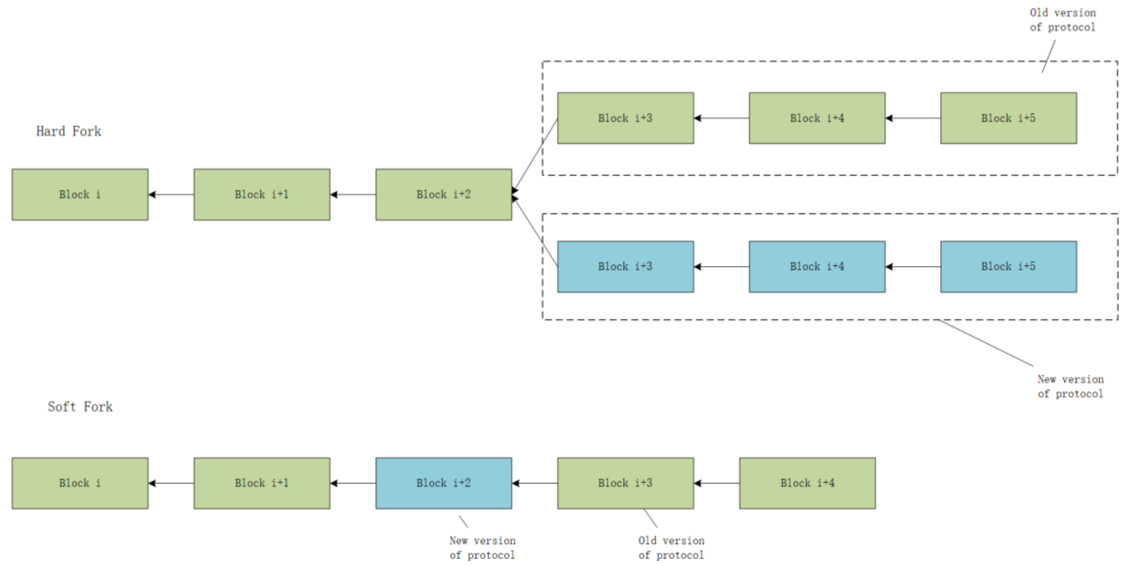


Figure 1. The difference between hard and soft forks
Note: Sort out and draw according to relevant imformation about blockchain forks

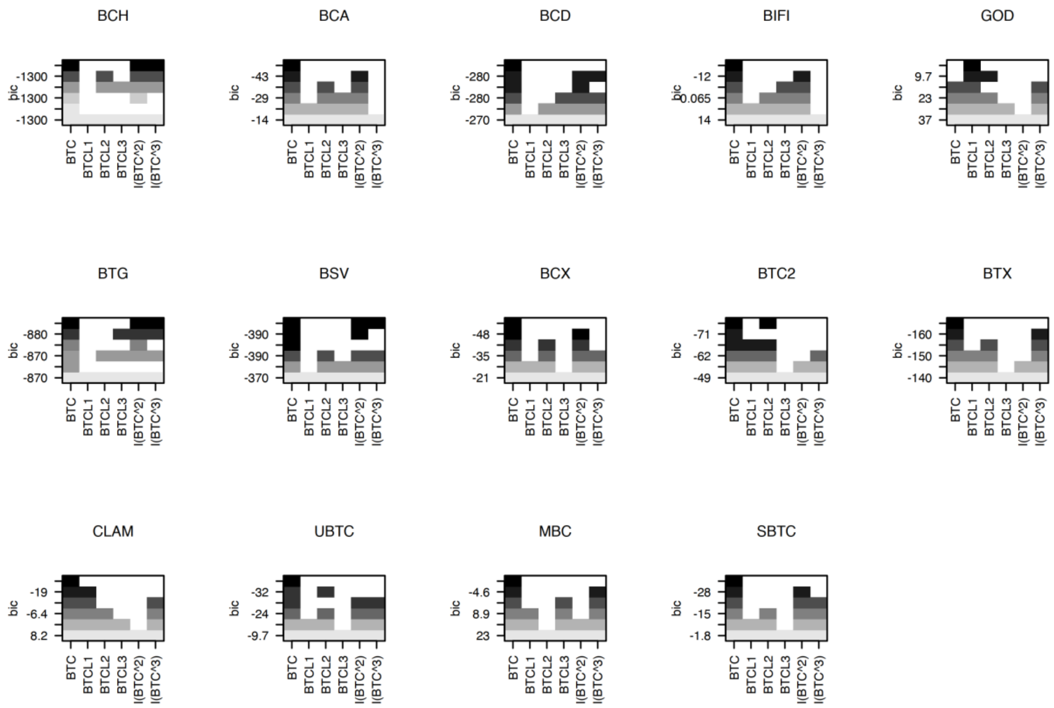
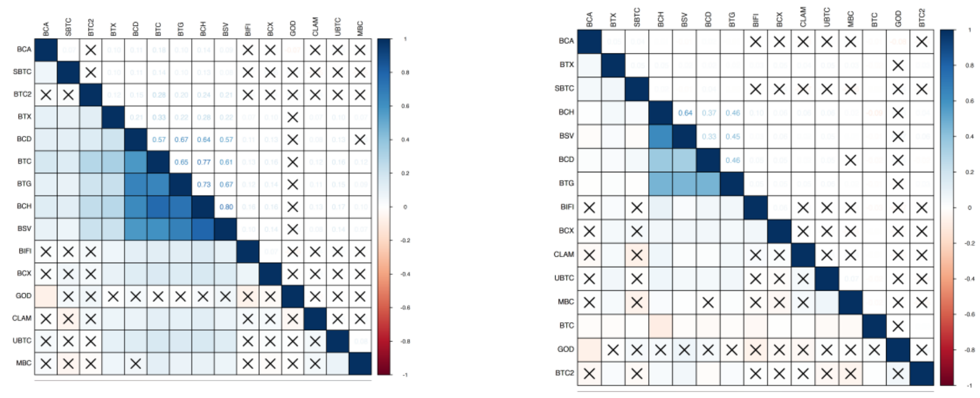


Figure 2. Variable selection based on BIC criterion



a b

Figure 3. Pearson correlation of BTC and its forked coins before and after filtering BTC influence

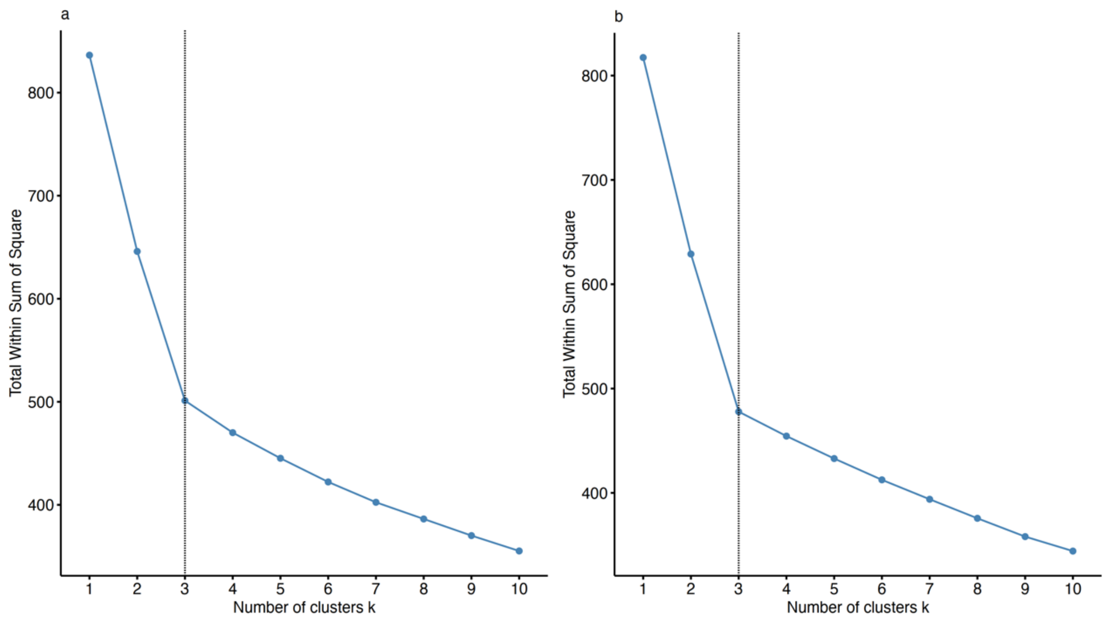


Figure 4. The number of clusters according to the elbow rule

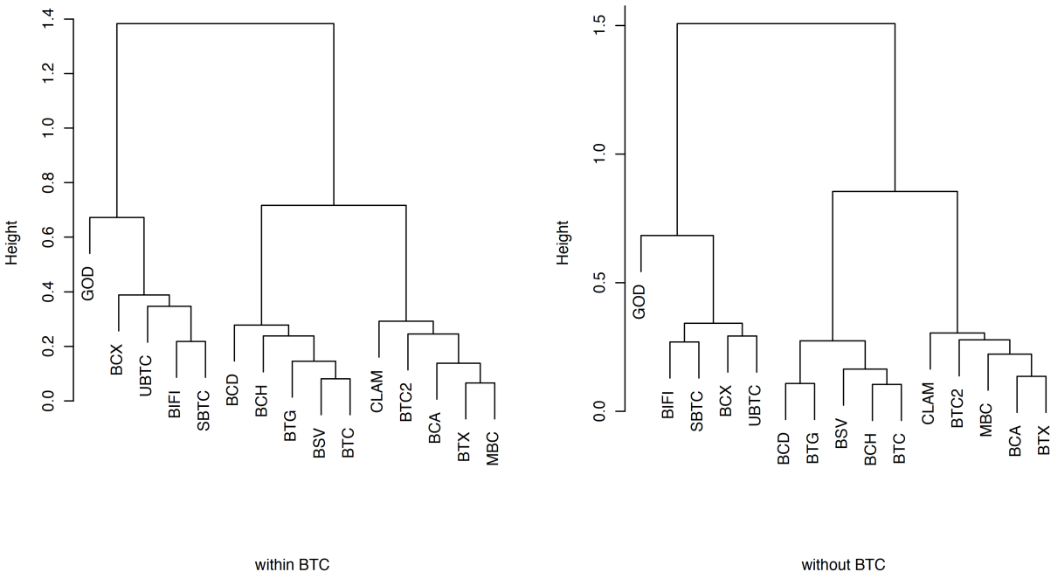


Figure 5. Cluster analysis before and after filtering BTC influence

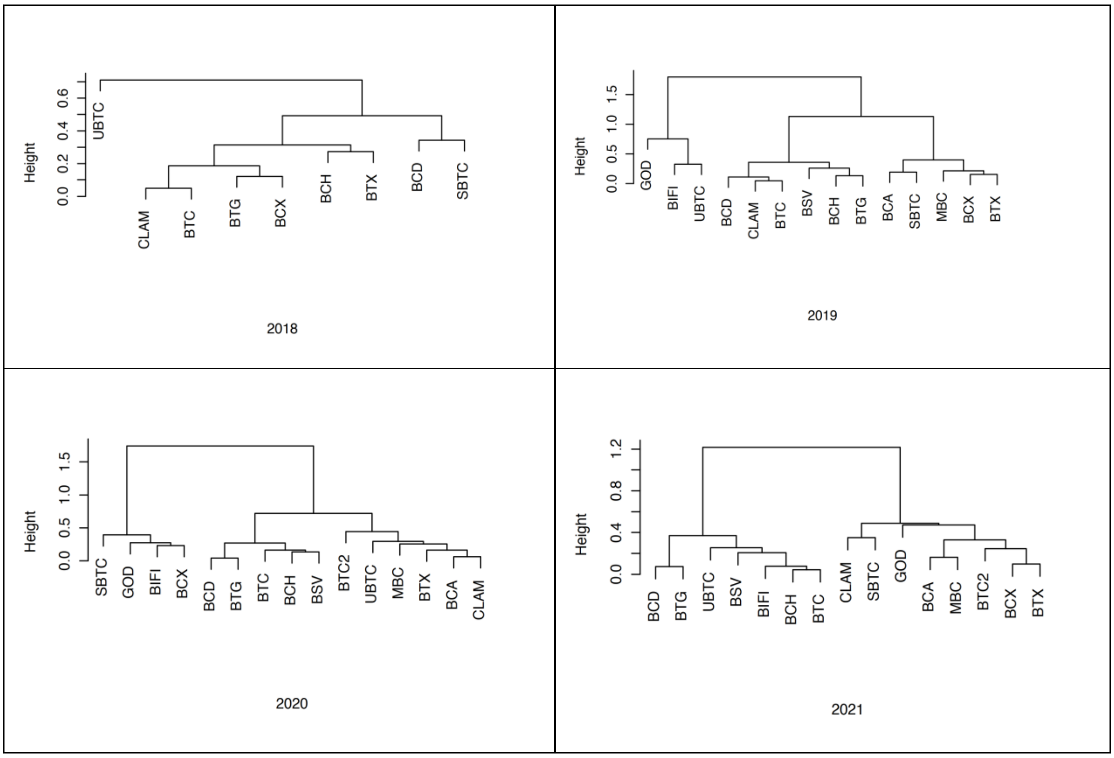


Figure 6. Cluster analysis by year after filtering out BTC influence