Physica A 509 (2018) 657–672

Contents lists available at ScienceDirect

Physica A

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

Quantifying the cross-correlations between online searches and Bitcoin market

Wei Zhang [a,b], Pengfei Wang [a], Xiao Li [c], Dehua Shen [a,b,*]

[a] College of Management and Economics, Tianjin University, Tianjin, 300072, PR China
[b] China Center for Social Computing and Analytics, Tianjin University, Tianjin 300072, PR China
[c] School of Finance, Nankai University, Tianjin, 300350, PR China

HIGHLIGHTS

- We quantify the cross-correlations between Change of Google Trends (CGT) and Bitcoin market with MF-DCCA.
- CGT-returns and CGT-volume are overall significantly cross-correlated.
- There exist power-law cross-correlations between CGT-returns and CGT-volume.
- The cross-correlations of CGT-returns have a higher degree of multifractal in the long-term and weak multifractal in the short-term.
- The cross-correlations scaling exponents are less than 0.5, which indicate anti-persistent cross-correlated.

ARTICLE INFO

Article history:
Received 8 February 2018
Received in revised form 31 March 2018
Available online 20 June 2018

Keywords:
Google Trends
Bitcoin
Cross-correlations
Multifractal detrended cross-correlation analysis
Rolling windows analysis

ABSTRACT

In this paper, we quantify the cross-correlations between Google Trends and Bitcoin market. By employing the Multifractal Detrended Cross-correlation Analysis (MF-DCCA) method, we find that the change of Google Trends (CGT) and Bitcoin market, i.e., returns and changes of volume, are overall significantly cross-correlated based on the cross-correlation test. In particular, the empirical results show that: (1) there exist power-law cross-correlations between CGT and Bitcoin returns as well as CGT and changes of volume; (2) the cross-correlations between CGT and returns have a higher degree of multifractal in the long-term and weak multifractal in the short-term, while the cross-correlations between CGT and change of volume show the opposite trend. (3) with the rolling window analysis, we further find that there is a decrease trend for the cross-correlations between CGT and Bitcoin returns over time, and the cross-correlations scaling exponents are less than 0.5, which indicate that they are both anti-persistent cross-correlated.

1. Introduction

With the development of information technology over the past two decades, the ways of generating, processing and transmitting information have undergone great changes and thus profoundly influence the asset prices in capital markets. Millions of searches, news, commentaries and recommendations are generated daily on social media, from which behavioral economics researchers extract proxies that reflecting investor sentiment. Online searches, for example, as one of online activities, have been deeply studied about its impact on the financial market, including predicting volatility [1], forecasting stock returns [2,3], quantifying trading behavior [4] and etc. Meanwhile, Bitcoin as one from of cryptocurrency has growth dramatically in recent years, which draw tremendous attention from online investors. Compared to the trading process for

* Correspondence to: College of Management and Economics, Tianjin University, No. 92 Weijin Road, Nankai District, Tianjin, 300072, PR China.
E-mail address: dhs@tju.edu.cn (D. Shen).

https://doi.org/10.1016/j.physa.2018.06.073
0378-4371/

other financial assets, e.g., stocks and bonds, the mining, information sharing and trading of cryptocurrency are completely inseparable from the Internet. In particular, search engines can retain the important traces of the generation, development and changes of emerging assets. Google Trends records search results for all search terms, and the public can download historical search index for certain keywords using Google Trends. Since the searches results are accurate and can be acquired directly, it makes possible to explore the cross-correlations between online searches and Bitcoin market. The feature of the limited amount of Bitcoin is significantly different from other assets, which makes it necessary to carry out this research.

As the leading cryptocurrency, Bitcoin's market capitalization once exceeded 305 billion U.S. dollars and surpassed large-scale financial companies such as VISA and Goldman Sachs.[1] Despite the widely concerns from the government, the media and the investors, the academic research on cryptocurrencies is limited due to its relatively short existence period. Current research focuses on a number of financial variables, including price formation [5,6] and the inefficiency of cryptocurrency [7–9], stylized facts [10], bubbles [11] and the relationship with other assets [12]. However, as a new thing using block chain technology in the Internet age, the Bitcoin's characteristics, especially its relationship with online activities, are not fully reflected in the existing financial research.

To supplement the conventional Pearson and Spearman correlation, many methods have been proposed by scholars. Podobnik and Stanley propose detrended cross-correlation analysis (DCCA) to study the power-law cross-correlation between two simultaneously recorded time series [13]. Podobnik et al. propose statistical tests for power-law cross-correlation [14]. By combining DCCA and multifractal detrended fluctuation analysis (MF-DFA) [15], Zhou introduced multifractal detrended cross-correlation analysis (MF-DCCA) [16] and it has been applied in many fields. The MF-DCCA method investigate cross-correlation taking into account the multifractal characteristics, yielding a more comprehensive analysis than the previous method. By employing the MF-DCCA method, this paper studies the cross-correlation between online searches and the financial variables of Bitcoin. An investigation of this correlation helps understand the impact of online activities on emerging assets.

The remainder of this paper is organized as follows. Second 2 is a brief literature review of online searches, Bitcoin and MF-DCCA method. Section 3 describes the data acquisition and processing. Section 4 introduces the details of the MF-DCCA method. Section 5 presents the empirical results and Section 6 concludes.

2. Literature review

2.1. Online searches and assets performance

Since search behavior is part of the online activity, let us review the literature on the latter. There are two clear lines of research literature. The first aspect is the correlation between investors' online activities and stock return performance. The main feature is: to use the content and quantity of posts in social platforms (e.g. Facebook [17], Twitter [18], microblogging [19], stock forum [20]) or search engine (e.g. Google [21,22], Baidu [3]) to construct proxy variables to study the relationship with stock market. The second aspect focuses mainly on the dynamics properties of the stock markets under the influence of online activities [23,24].

With regard to online searches, the existing literature also focuses on two aspects. The first is to measure whether the Search Volume Index (henceforth SVI) can predict financial variables to explain the anomalies [25]. Da et al. proposed a new proxy of investor attention using Google search frequency and found that the proxy was able to capture investor attention in a more timely fashion, to measure the retail investor's attention, and to predict the price of the next two weeks [26]. Vozlyublennaia found that with the increase of SVI, returns predictability would be diminished [27]. The second is to study the relationship between online searches and stock market from the perspective of information needs and dissemination. Zhang et al. proved that Internet enhanced the speed of information diffusion [3]. And Drake et al. found that SVI represented the information demand and spiked at the announcement [28].

2.2. Bitcoin

The existing literature on Bitcoin mainly focuses on the inefficiency, stylized facts and relationship with other assets. Kristoufek utilized the Efficiency Index and found that Bitcoin markets denominated in USD and CNY remain mostly inefficient between 2010 and 2017 [9]. Bariviera employed the rolling window Detrended Fluctuation Analysis method and found that Bitcoin was in the process of moving to efficient and the volatility of Bitcoin had long-term memory throughout the sample period [29]. Alvarez-Ramirez et al. estimated long-range correlations for price returns and found Bitcoin market presented asymmetric correlations with respect to increasing and decreasing price trending [30]. Baur et al. found no correlation between Bitcoin and other traditional assets, either in the normal period or the financial crisis, and drew conclusion from the detailed trading data that the main use of Bitcoin was speculation [31]. Garcia et al. estimated the relation between bubbles of Bitcoin and social interaction from the perspective of price, volume of communication in online media, volume of searches and user growth [11]. There are also some studies on Bitcoin banning, transaction costs and price manipulation [32–34]. Kristoufek used Bitcoin data with a duration of no more than two years to study the relation

[1] https://cointelegraph.com/news/bitcoins-price-surpasses-18000-level-market-cap-now-higher-than-visas (assessed on February 1, 2018)

Table 1
Basic information of each USD-rate Bitcoin market.

Market	Start date	Average volume(BTC)
Abucoins	2017-9-21	9.6588
Bitbay	2014-5-16	14.4285
Bitkonan	2013-7-2	2.1770
Bitstamp	2011-9-13	12873.2950
Btcalpha	2016-11-1	147.6091
Btcc	2016-11-2	87.7846
Cex	2014-7-18	1170.6564
Coinbase	2014-12-1	15057.8374
Coinsbank	2017-7-3	1701.1693
Getbtc	2017-6-16	851.6766
Itbit	2013-8-25	2785.9914
Kraken	2014-1-7	5048.1215
Lake	2014-3-1	4872.0768
Localbtc	2013-3-11	529.2942
Okcoin	2017-8-7	559.4005
Wex	2017-9-19	1815.5089

between search queries and the prices, and shows that there also existed a pronounced asymmetry between the effect of an increased interest in the currency while being above or below its trend value [35]. Yelowitz and Wilson conducted four proxies including speculative investors and criminals from Google Trends to analyze the behaviors of Bitcoin's investors [36]. However, it is necessary to estimate the cross-correlation between Google trends and Bitcoin market that covers a longer time span.

2.3. Multifractal detrended cross-correlation analysis

The existing literature can be divided into two parts. The first part is to propose and improve the method. DFA, DMA [37], DCCA [13] and their variants that have been popularly applied in the field of econophysics [38–42] promote the emergence of MF-DFA [15] and MF-DMA [43,44]. Base on the MF-DFA and DCCA, Zhou first introduced the multifractal detrended cross-correlation analysis [16]. Then a series of improvement methods have been proposed, such as the method of MF-X-DMA [45], MF-HXA [46], MF-SMXA [47], MF-X-PF [47–49], MFCCA [50,51] MF-PX-DFA [52], MF-PX-DMA [52], MF-X-WT [53], MF-X-WL [54] and etc. And Yue et al. designed an objective procedure to determine the two Hurst indexes delimited by the crossover scale [55]. These are the applications of MF-DCCA in some cases and improve the efficiency of the method. The second part is the application of MF-DCCA to find some empirical results. Ma et al. investigated the cross-correlations between the crude oil and the six GCC stock markets and employed rolling window method to investigate the dynamic of the scaling exponent [56]. Zhuang et al. studied the nonlinear structure of the cross-correlations between carbon and crude oil markets based on MF-DCCA [57].

3. Data description

The sample data we use in this paper have two sources, the first is Bitcoin data for different trading platforms crawled from https://bitcoincharts.com/, and the second is Google Trends directly download from https://trends.google.com/. The Bitcoincharts website classifies all Bitcoin exchanges as active and inactive. The data obtained from Bitcoincharts website include open, high, low, close, volume (BTC), volume (currency) and weighted price of all active Bitcoin market. Bitcoin exchange rate displayed on Bitcoinchats are United States Dollar (USD), Euro, Japanese Yen, Korean Won, British Pound, Australian Dollar and etc. Due to the different time of the establishment of the platform and the start of the trading, the time span of the Bitcoin data varies. In order to choose a currency rate that can adequately represent global Bitcoin trading, we sort the trading volume of each currency rate on December 31, 2017. The Japanese Yen rate and the USD rate were the top two. Total trading volume of six Japanese-yen-rate markets is 99797 (BTC), and 16 USD-rate markets is 40152 (BTC). The earliest yen rate Btcbox market could be traded on Apr 9, 2014. Because the yen data only covers a part of the timespan of Bitcoin, we choose the USD data that can be traced back to 2011. In order to avoid trading platform bias, we generate a daily-volume-weighted composite index of USD-rate Bitcoin. Table 1 reports the basic information of 16 USD-rate Bitcoin market, and the average volume denotes the average daily trading volume (BTC) of each market in 2017. In the synthesis procedure, we weighted sum the closing price that are denominated in USD-rate to get index price

$$IndexPrice_t = \sum_{i=1}^{n} cp_{i,t} * \frac{Volume_{i,t}}{TotalVolume_t} \tag{1}$$

where $cp_{i,t}$ and $Volume_{i,t}$ denote closing price and market value of i_{th} symbol at t day, and $TotalVolume_t = \sum_{i=1}^{n} Volume_{i,t}$. The volume of composite index is the sum of symbols denominated in USD-rate.

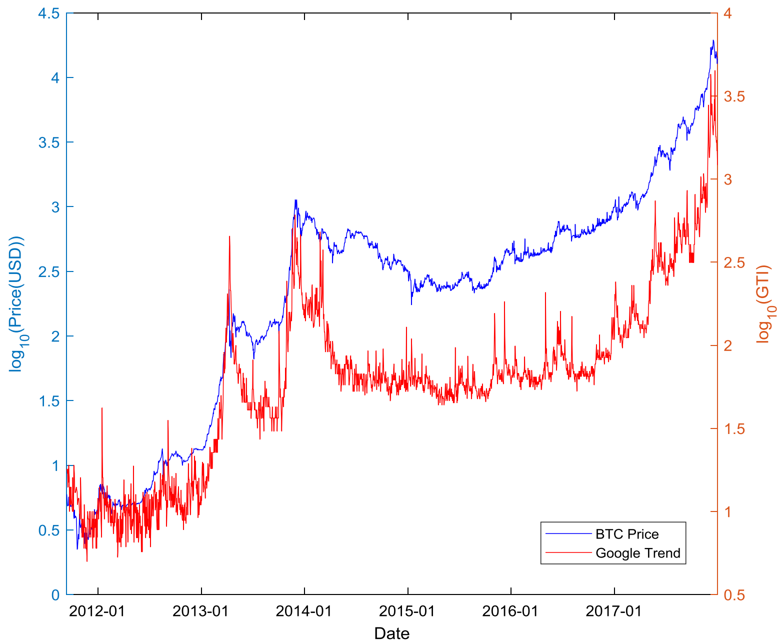

Fig. 1. Bitcoin index and Google trends index.

We search the term "Bitcoin" and directly download the original data. As we all know, Google Trends returns data of different frequencies based on the length of time chosen. If the time span is more than five years, we can only obtain monthly data. In order to get accurate daily data, we download Google Trends by time segment, with a two-month overlap period for each segment. We calculate the cumulative multiplier of the overlap segment, and then multiply the multiplier by the original data of the time segment to obtain a composite Google Trends Index (henceforth GTI). The GTI data from June 1, 2011 to February 1, 2017 are the same as the Google Trends downloaded directly from the website. Fig. 1 shows GTI and composite index of USD-rate Bitcoin price, showing a strong positive correlation. Fig. 2 shows the trading volume of Bitcoin index. As is clearly shown in Fig. 2, daily trading volume does not exceed 200,000 Bitcoins.

In order to measure the cross-correlation between GTI and Bitcoin index subsequently, we need to define three variables, i.e. returns, changes of volume and changes of GTI with equation:

$$Ret_t = ln(\frac{p_t}{p_{t-1}}) \tag{2}$$

$$CV_t = \ln\left(\frac{volume_t}{volume_{t-1}}\right) \tag{3}$$

$$CGT_t = ln(\frac{GTI_t}{GTI_{t-1}}) \tag{4}$$

where p_t, $volume_t$ and GTI_t indicate the index price, trading volume and GTI on trading day t respectively. Table 2 shows the statistical properties of Bitcoin index and GTI. The standard deviation of CV is the largest, showing that change of volume is the most volatile. The Jarque–Bera statistics of three variables indicate that they all reject the null hypothesis that time series come from a normal distribution with an unknown mean and variance. And the kurtosis that are greater than 3 show the distributions of three time series are more outlier-prone than the normal distribution. Henceforth, Ret, CV and CGT denote returns, change of Bitcoin and change of Google Trends Index.

4. Methodology

We employ the popular MF-DCCA method proposed by Zhou [16] to analyze the cross-correlation of two nonstationary series. Let us briefly introduce this method. Consider two time series $\{x_i\}$ and $\{y_i\}$ of the same length N, where $i = 1, 2, \ldots, N$. Then the MF-DCCA method can be described as follow:

Step 1. The profile is determined

$$X_i = \sum_{k=1}^{i}(x_k - \bar{x}), \qquad Y_i = \sum_{k=1}^{i}(y_k - \bar{y}), \quad i = 1, 2, \ldots, N \tag{5}$$

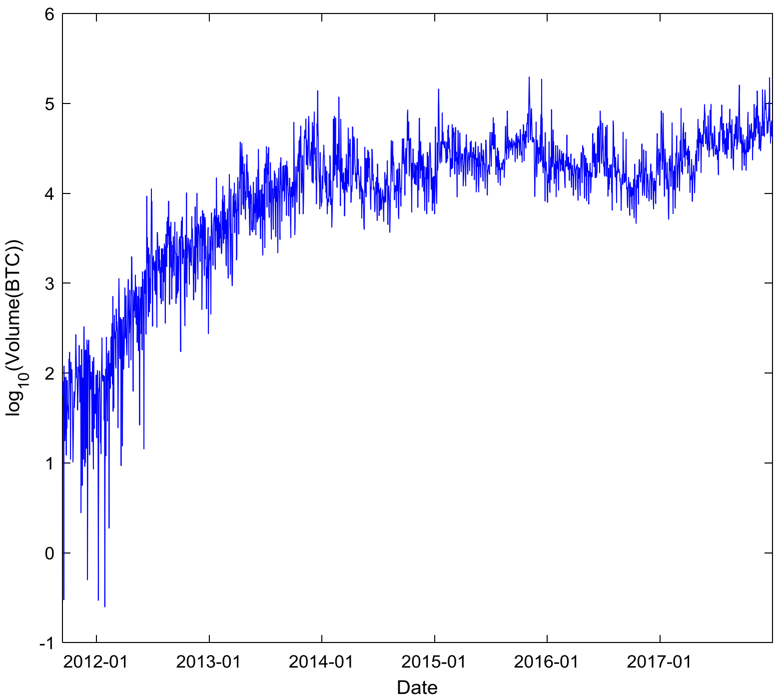

Fig. 2. Volume of Bitcoin index.

Table 2
Statistical properties of Bitcoin index and Google Trend Index.

	Obs	Mean	Median	Std	Max	Min	Skewness	Kurtosis	Jarque–Bera
Ret	2283	0.0034	0.0038	0.0590	0.4455	−0.6636	−1.0334	20.0193	27959.7828***
CV	2283	0.0029	−0.0289	0.7609	6.0276	−5.9090	0.1324	13.5008	10495.8491***
CGT	2283	0.0018	0.0000	0.1738	1.3398	−0.8675	0.8178	8.8678	3529.7076***

***Indicates statistical significance at 1% level.

where $\overline{x}$ and $\overline{y}$ denote the average of $\{x_i\}$ and $\{y_i\}$.

Step 2. The integrated time series X_i and Y_i can be used to analyze long-range correlations in $\{x_i\}$ and $\{y_i\}$ readily by getting rid of trends. Then we divide the two profiles into $N_s \equiv int(N/s)$ non overlapping segments of equal length s respectively. If the length N is not an integer multiple of scale s, there will remain a short part segment at the end of each profile. In order not to neglect a short part at the end of the profile, we repeated the same divided procedure starting from the end of the profile, so $2N_s$ segments are obtained altogether. The N_s is set as $10 < N_s < N/4$ and the step is 1 in this paper.

Step 3. The local trend of the series $\tilde{X}_v(i)$ and $\tilde{Y}_v(i)$ is determined for each segment v

$$\tilde{X}_v\,(i) = \tilde{a}_k i^m + \cdots + \tilde{a}_1 i + \tilde{a}_0 \quad \tilde{Y}_v\,(i) = \tilde{b}_k i^m + \cdots + \tilde{b}_1 i + \tilde{b}_0 \tag{6}$$

where $i = 1, 2, \ldots, s$, $v = 1, 2, \ldots, 2N_s$, $m = 1, 2, \ldots$.

Then we calculate the polynomial fit of each segments v and obtain the detrended covariance

$$F^2\,(s, v) = \frac{1}{s}\sum_{j=1}^{s}\left[X((v-1)\,s+j) - \tilde{X}_v(i)\right]\left[Y((v-1)\,s+j) - \tilde{Y}_v(i)\right] \tag{7}$$

for $v = 1, 2, \ldots, N_s$, and

$$F^2\,(s, v) = \frac{1}{s}\sum_{j=1}^{s}\left[X(N-(v-N_s)\,s+j) - \tilde{X}_v(i)\right]\left[Y(N-(v-N_s)\,s+j) - \tilde{Y}_v(i)\right] \tag{8}$$

for $\upsilon = N_s + 1, N_s + 2, \ldots, 2N_s$, and where $\tilde{X}_\upsilon(i)$ and $\tilde{Y}_\upsilon(i)$ is the local trend.

Step 4. Get the qth-order fluctuation function by averaging over all segments,

$$F_q(s) = \{\frac{1}{2N_s}\sum_{\lambda=1}^{2N_s} F^2(s,\upsilon)^{q/2}\}^{1/q} \tag{9}$$

And when $q = 0$, the equation is defined by

$$F_0(s) = \exp\{\frac{1}{4N_s}\sum_{\lambda=1}^{2N_s} \ln\left[F^2(s,\upsilon)\right]\} \tag{10}$$

Step 5. Analyze the scale behavior of the fluctuation function by observing the log–log plots of $F_q(s)$ versus s. If two series are long-range cross-correlated, we obtain a power-law relationship as follows,

$$F_q(s) \propto s^{H_{xy}(q)} \tag{11}$$

where $H_{xy}(q)$ is the well-known Generalized Hurst exponent.

Step 6. The Renyi exponent $\tau_{xy}(q)$ can be used to characterize the multifractal nature,

$$\tau_{xy}(q) = qH_{xy}(q) - 1 \tag{12}$$

If the Renyi exponent function is a linear function of q, the cross-correlation of the two correlated series is monofractal; otherwise, it is multifractal.

The singularity spectrum $f(\alpha)$ that describes the singularity content of the time series can be obtained through Legendre transform:

$$\alpha_{xy}(q) = H_{xy}(q) + H'_{xy}(q) \tag{13}$$

$$f_{xy}(\alpha) = q\left[\alpha_{xy} - H_{xy}(q)\right] + 1 \tag{14}$$

Here, α is the holder exponent that characterizes the singularities in time series.

5. Empirical results

5.1. Cross-correlation test

Scholars have proposed many method to measure power-law cross-correlation [58–61]. Following existing literature, we conduct a preliminary analysis by employing cross-correlation test introduced by Podobnik et al. [59] in analogy to Ljung–Box test [62]. Ma et al. [63] used this approach to investigate cross-correlations between several stock markets. Wang et al. [64] also applied the method to test the cross-correlations A-share and B-share in Chinese stock market. For two time series $\{x_i\}$ and $\{y_i\}$ with the equal length N, the test statistic is defined as

$$Q_{cc}(m) = N^2\sum_{i=1}^{m}\frac{X_i^2}{N-i} \tag{15}$$

where X_i^2 is the cross-correlation function

$$X_i = \frac{\sum_{k=i+1}^{N} x_k y_{k-i}}{\sqrt{\sum_{k=1}^{N} x_k^2 \sum_{k=1}^{N} y_k^2}} \tag{16}$$

The test statistic $Q_{cc}(m)$ is approximately $\chi^2(m)$ distributed with m degrees of freedom. We can use the statistic to test the null hypothesis that none of the first m cross-correlation coefficient is different from zero. Thereby, if the value of Q_{cc} is larger than the critical value of $\chi^2(m)$, the cross-correlation between the two time series is significant.

Fig. 3 shows the cross-correlation statistics $Q_{cc}(m)$ for two pairs of time series. The black line is the critical value of $\chi^2(m)$ at 5% significant level, and the degrees of freedom vary from 1 to 1000. Regardless of m changes, $Q_{cc}(m)$ of the blue line with asterisk is always greater than the value of critical value of $\chi^2(m)$. For $log\,10(m) > 2.5$, $Q_{cc}(m)$ of the red line with circle is larger than critical value and has the trend of being larger than $\chi^2(m)$. Thus, we can reject the null hypothesis of no cross-correlations.

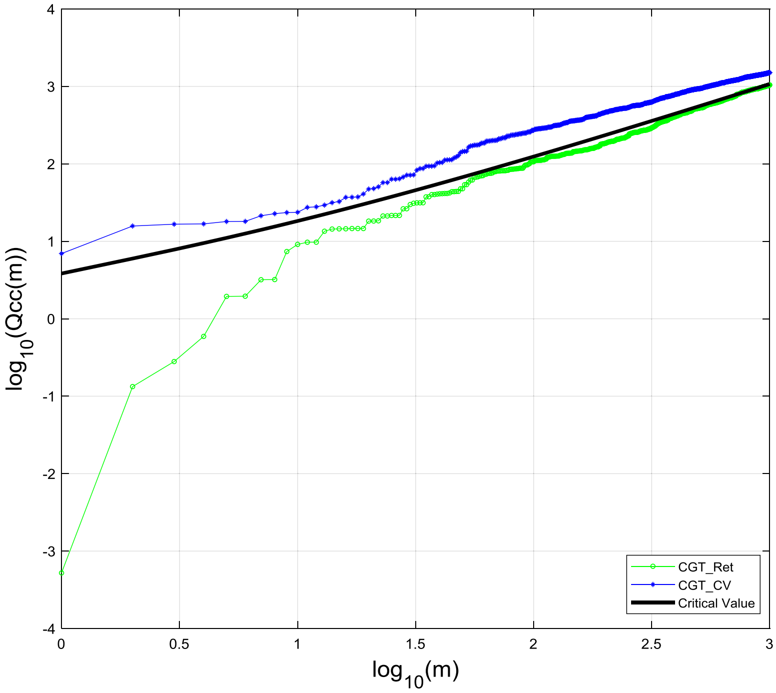

Fig. 3. Cross-correlation test (Q_{cc} (m)).

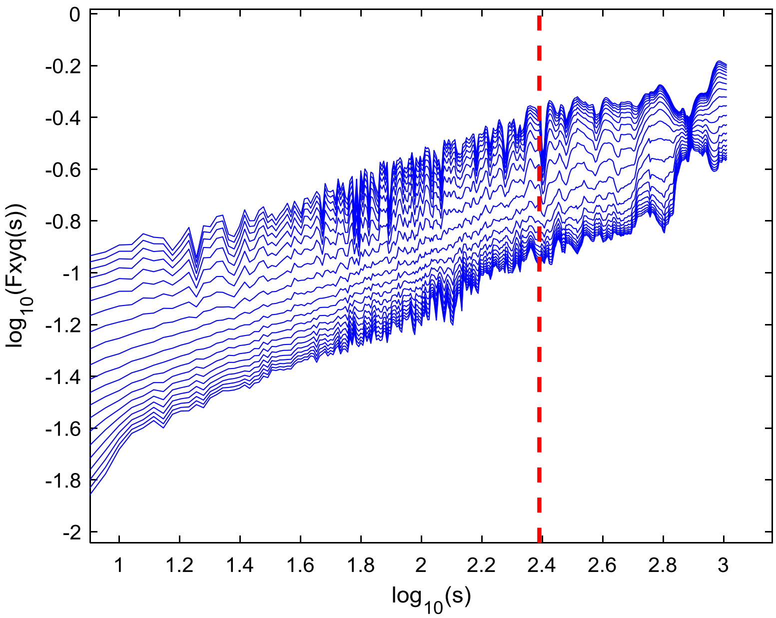

Fig. 4. Log–log plots of $F_{xyq}(s)$ versus s for *CGT* and *Ret*.

5.2. Multifractal detrended cross-correlation analysis

Since the value of Q_{cc} (m) has qualitatively demonstrated the existence of cross-correlation in Section 5.1, we next employ MF-DCCA quantitatively to study the cross-correlation of *CGT* and *Ret* as well as *CGT* and *CV*. In this papers, the scale q is set from -10 to 10 and the step size is set to 1. Fig. 4 and Fig. 5 are the Log–log plots of fluctuation function $F_{xyq}(s)$ versus time scale s for *CGT* and *Ret*, as well as *CGT* and *CV* respectively. As we see in Fig. 4 and Fig. 5, all lines fit the logarithmic line of $F_{xyq}(s)$ versus s for q from -10 to 10, which provide evidence of power-law cross-correlation between the two pairs.

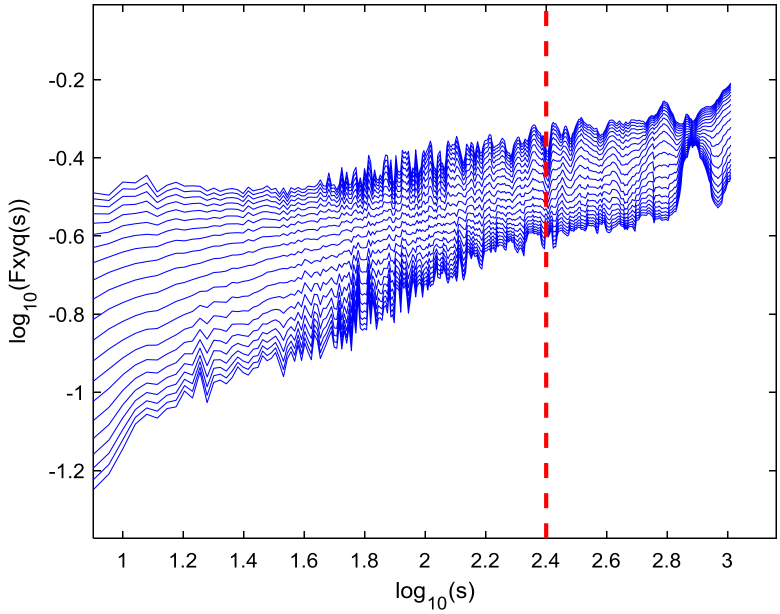

Fig. 5. Log–log plots of $F_{xyq}(s)$ versus s for *CGT* and *CV*.

Table 3
Cross-correlation scaling exponents for two pairs with q varying from −10 to 10.

q	CGT_Ret		CGT_CV	
	$S^* = 252$		$S^* = 246$	
	$S < S^*$	$S > S^*$	$S < S^*$	$S > S^*$
−10	0.5271	0.7115	0.3890	0.2900
−9	0.5190	0.7119	0.3787	0.2894
−8	0.5097	0.7132	0.3667	0.2892
−7	0.4993	0.7157	0.3527	0.2895
−6	0.4879	0.7199	0.3365	0.2905
−5	0.4761	0.7260	0.3182	0.2923
−4	0.4646	0.7343	0.2978	0.2950
−3	0.4544	0.7431	0.2758	0.2980
−2	0.4468	0.7474	0.2526	0.2998
−1	0.4431	0.7335	0.2291	0.2976
0	0.4453	0.6812	0.2065	0.2892
1	0.4531	0.5864	0.1863	0.2740
2	**0.4610**	**0.4796**	**0.1690**	**0.2544**
3	0.4615	0.3920	0.1547	0.2336
4	0.4529	0.3301	0.1428	0.2139
5	0.4384	0.2879	0.1326	0.1967
6	0.4221	0.2585	0.1237	0.1822
7	0.4062	0.2373	0.1159	0.1702
8	0.3919	0.2215	0.1088	0.1603
9	0.3794	0.2092	0.1025	0.1522
10	0.3687	0.1993	0.0968	0.1454
ΔH_q	0.1584	0.5480	0.2922	0.1544

We define "crossover" s^* as a turning point when the linear tendency of the curve changes fundamentally. Short-term financial market behavior is vulnerable to external events of the market, while long-term behavior is determined by internal factors. Over time, the short-term shocks will gradually decay, and the long-term market supply and demand mechanism will exert its influence. The scale exponent of $s < s^*$ reflects short-term correlation, which also means short-term related behavior. And scale exponent of $s > s^*$ means long-term related behavior. Therefore, it can be said that "crossover" can reflect the duration of the impact of the factors that determine the short-term market behavior. According to the log–log plots Figs. 4 and 5, we can find the "crossover" of *CGT* and *Ret* located at approximately $\log(s^*) = 2.401$ (252 days, about 8.4 months), the "crossover" of *CGT* and *CV* at about $\log(s^*) = 2.39$ (246 days, about 8.2 months).

To study the different long-term and short-term changes in the cross-correlation, we calculate the cross-correlation scaling exponents for q from −10 to 10. The results are shown in Table 3. H_{xy} (2) of short term is smaller than that of long term for both CGT_Ret and CGT_CV, indicating the cross-correlations of short term are more anti-persistent. Here ΔH_q is

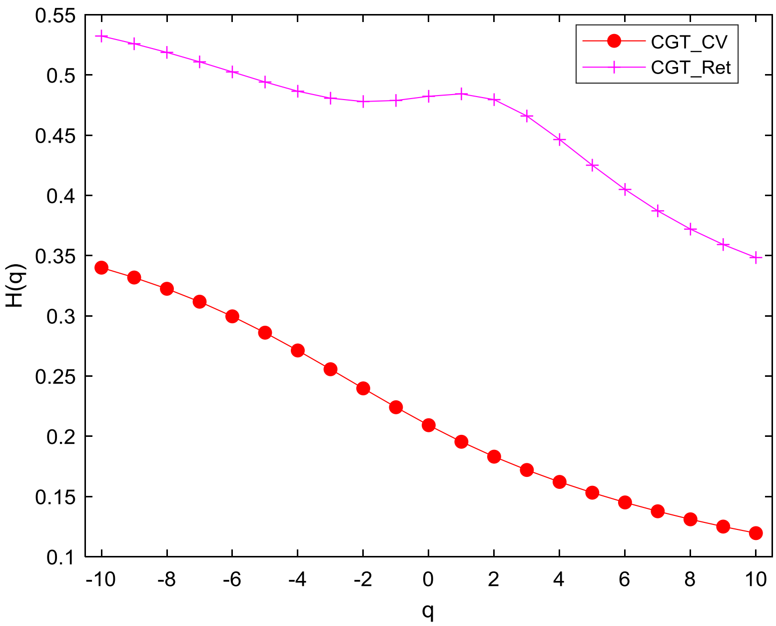

Fig. 6. Cross-correlation Hurst exponent, $H_{xy}(q)$ as a function of q for *CGT* and Bitcoin.

introduced to explore the degree of multifractal. The larger ΔH_q, the higher the degree of multifractality.

$$\Delta H_q = \max\left(H_q\right) - \min(H_q) \tag{17}$$

Before analyzing long-term and short-term characteristics, let us observe the overall cross-correlation characteristics. When q = 2, MF-DCCA is identical to DCCA, and the bivariate Hurst exponent H_{xy} (q) has similar properties and interpretation to the univariate Hurst exponent. Fig. 6 shows the relationship between the scaling exponents and q. H_{xy} (2) of cross-correlation between *CGT* and *Ret* as well as *CGT* and *CV* are 0.4795, 0.1831 respectively (also shown in Fig. 16), which indicate both of them are anti-persistently cross-correlated. H_{xy} (2) of *CGT* and *Ret* is lower than but very close to 0.5, indicating that the anti-persistency of cross-correlation is weakly. However, H_{xy} (2) of *CGT* and *CV* is far less than 0.5, showing that the anti-persistency of cross-correlation is strong. The difference of H_{xy} (2) of *CGT_CV* and *CGT_RET* may be caused by the volatile degree of *Ret* and *CV* series. The variance of the exponent and q indicates a cross-correlation multifractality between two pairs. Especially for the cross-correlation between *CGT* and *Ret*, the scale exponent for q < 0 is much larger than the s exponent for q > 0. Therefore, we draw conclusion that the small-fluctuation is more cross-correlated than the large one.

Figs. 7 and 8 present the scaling exponents for *CGT* and *Ret* as well as *CGT* and *CV* with q varying from -10 to 10. For the cross-correlation of *CGT* and *Ret*,when $s > s^*$, the scaling exponents decrease from 0.71 to 0.20, indicating that the cross-correlation is multifractal in the long-term. In addition, when $q < 0$, the exponents is greater than 0.5, indicating that the cross-correlation of small fluctuations in the long-term persists (positive). When $q > 2$, the exponents is less than 0.5, indicating that the cross-correlation behavior is anti-persistent in the long-term. However, for $s < s^*$, the value of $H_{xy}(q)$ is weakly dependent on q, indicating that short-term cross-correlation is weakly multifractal. In addition, almost all values of $H_{xy}(q)$ are less than 0.5, indicating short-term cross-correlation behavior is not persistent. And cross-correlation of *CGT* and *CV* is the opposite, compared to the former. Both in long-term and short-term scaling exponents are less than 0.5, indicating an anti-persistent cross-correlation.

We calculate the Renyi exponent $\tau_{xy}(q)$ that characterize the multifractality based on Eq. (11). According to Fig. 9, $\tau_{xy}(q)$ is a nonlinear curve of *CGT* and *Ret* as well as *CGT* and *CV* dependent on q. From Fig. 10 and Fig. 11, We find the *CGT* and *Ret* show a multifractal cross-correlation in the long term, and $\tau_{xy}(q)$ is nearly linear in the short term, which indicates a nearly monofractal cross-correlation in the short term. For *CGT* and *CV*, the conclusion is exactly the opposite, and the degree of multifractality below the former. This evidence supports the findings of Fig. 7, Fig. 8 and Table 3.

Then we use Eqs. (13)–(14) to calculate the multifractal spectrum of *CGT* and *Ret* as well as *CGT* and *CV* respectively. The broad spectrum implies strong multifractality and narrower spectrum indicates the weak multifractal behavior. If the multifractal spectrum appears as a point, it is monofractal. The multifractal spectrum of *CGT* and *Ret* as well as *CGT* and *CV* is not a key point, there are the multifractal characteristics between the two pairs. In Figs. 12–14, although some points are clustered together, the width of the spectrum is not zero. As can be seen from Fig. 13, the multifractal strength of spectrum of *CGT* and *Ret* in long-term is stronger than that in short-term. And the spectrum of *CGT* and *CV* is just the opposite.

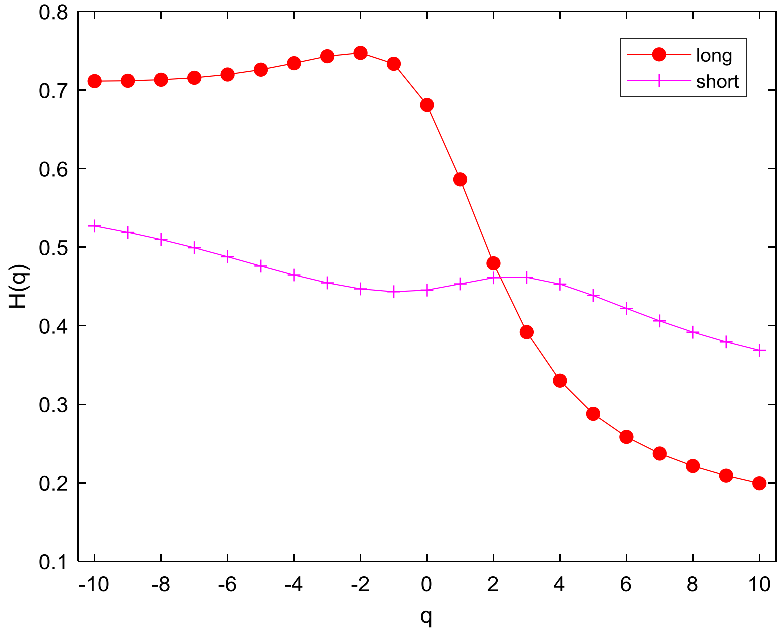

Fig. 7. Scaling exponents for *CGT* and *Ret* with q varying from −10 to 10.

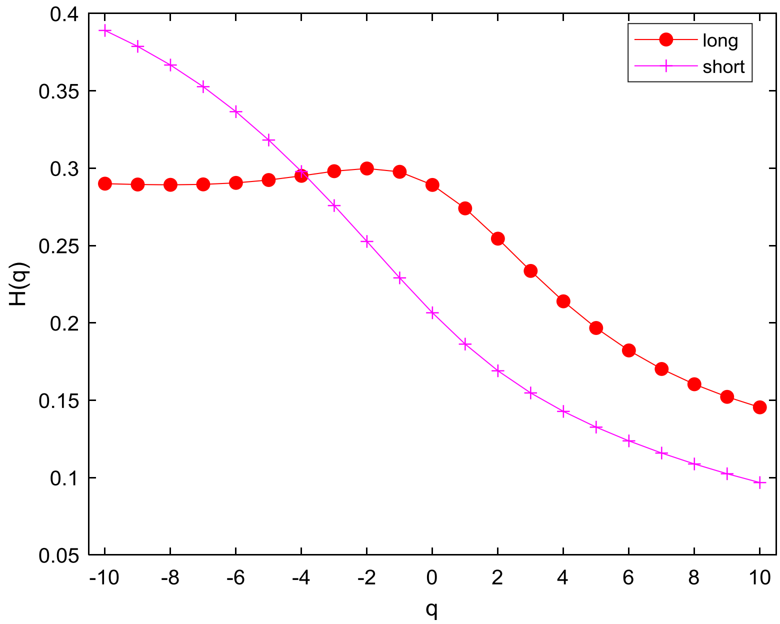

Fig. 8. Scaling exponents for *CGT* and *CV* with q varying from −10 to 10.

5.3. Rolling window analysis

With the ability to dynamically evaluate the informational efficiency over time, the rolling window approach has been widely used to investigate many topics in financial markets. In this section, we employ the rolling window method of MF-DCCA to investigate the dynamics of the cross-correlation.

According Bariviera [10,29], the length of windows chosen in this paper is 500 days. And we use six points to estimate the scaling exponent, the points for regression estimation are: $s = \{8, 16, 32, 64, 128, 256\}$. Fig. 15 shows the time varying exponents H_{xy} (2). The red line is the exponent for the cross-correlation between *CGT* and *Ret*, and the green is for the one between *CGT* and *CV*. Although the two curves are volatile, the trend is still clear. We can see that on the whole, the scaling exponents are on a downward trend. Specifically, before July 2014, the two time varying scaling exponents are in the

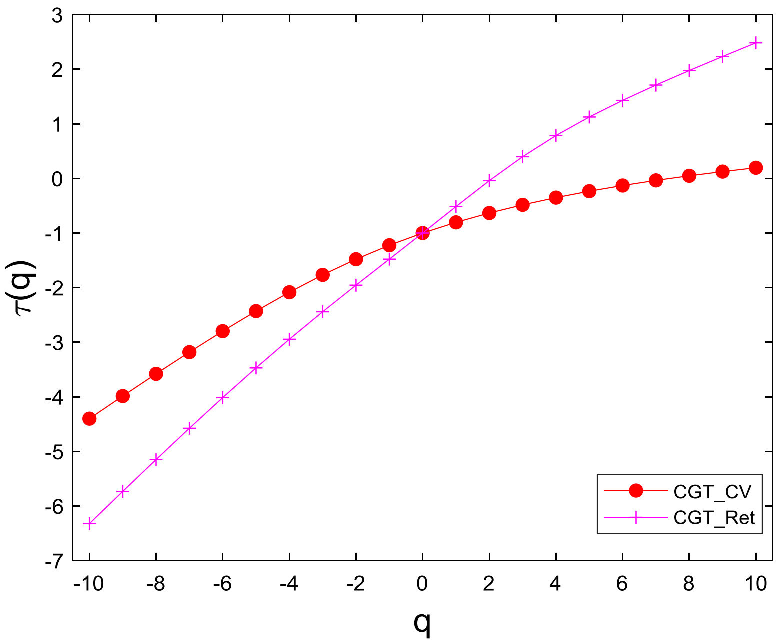

Fig. 9. The relationship $\tau_{xy}(q)$ and q for *CGT* and Bitcoin.

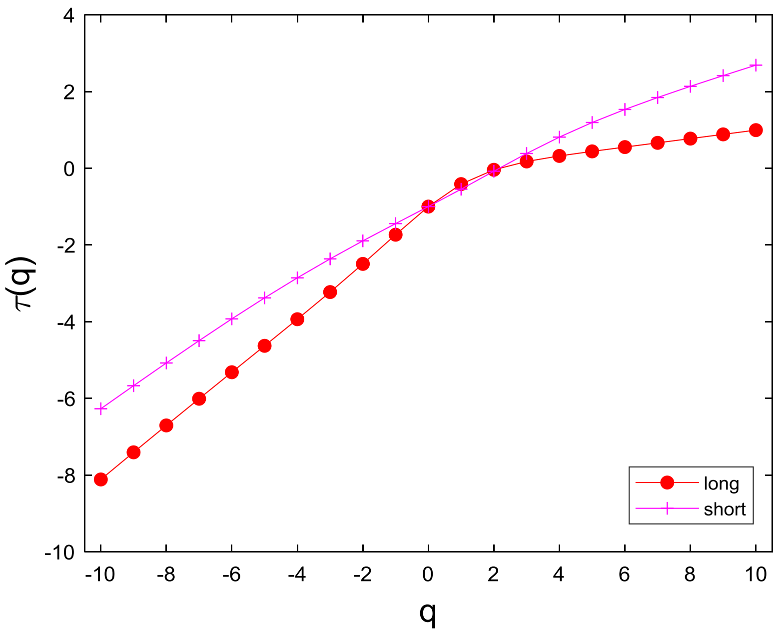

Fig. 10. The relationship $\tau_{xy}(q)$ and q for *CGT* and *Ret* both in long-term and short-term.

process of increase. However, they subsequently experience another year of decline before beginning to fluctuate around a certain value after July 2015. Now, both scaling exponents are below 0.5, meaning that volatilities of *CGT* are not persistently cross-correlated. In addition, we also find that the exponent for the *CGT* and *Ret* is larger than *CGT* and *CV* in the entire sample period. This indicates that time series of *CGT* and *CV* are more anti-persistently cross-correlated. We speculate that the relatively low value of H_{xy} (2) for *CGT* and *CV* may be related to the limited amount of Bitcoin. Even if bitcoin miners continue mining, the number of Bitcoins will not exceed the initial limit of the design. Although the speculative boom has increased the frequency of trading, due to the total amount of limitation, the daily trading volume will not increase by orders of magnitude. We reason that the low value of H_{xy} (2) for *CGT* and *CV* may be due to the fact that the daily trading volume

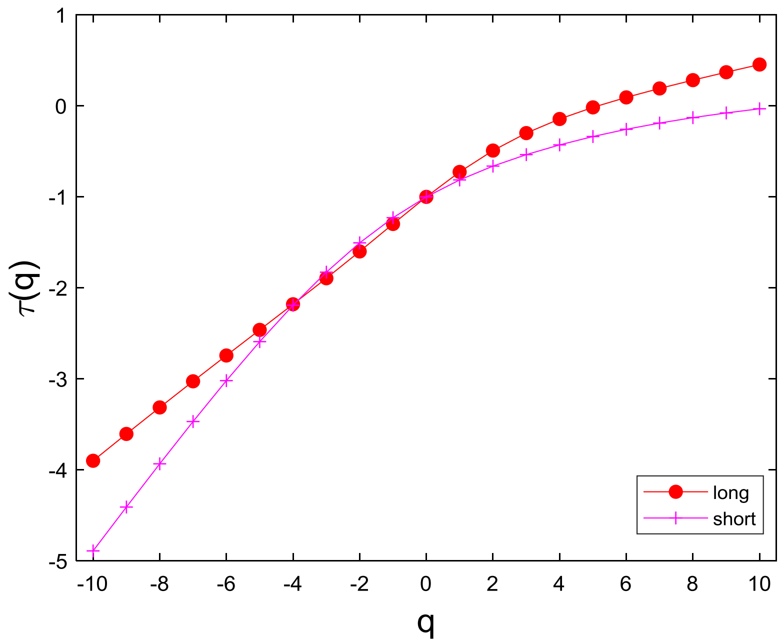

Fig. 11. The relationship $\tau_{xy}(q)$ and q for *CGT* and *CV* both in long-term and short-term.

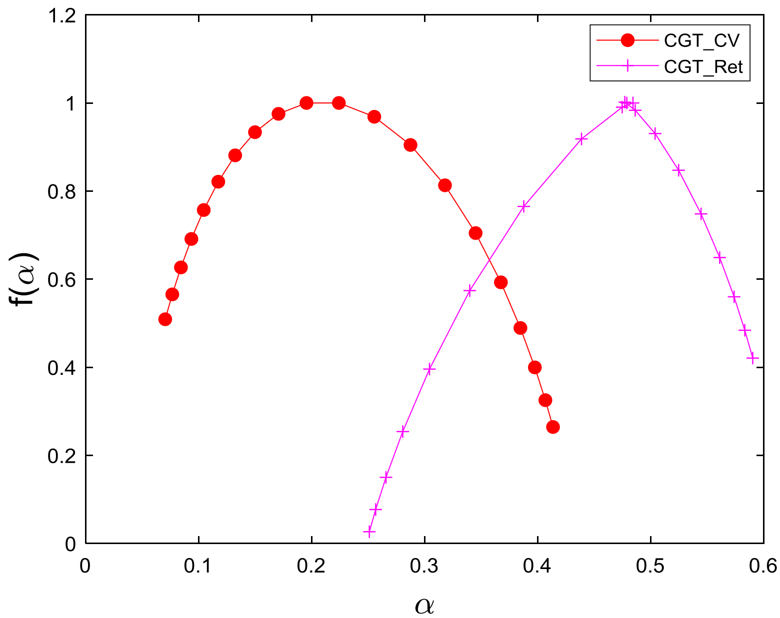

Fig. 12. The multifractal spectrum, $f_{xy}\,(\alpha)$ as a function of α for *CGT* and Bitcoin.

(BTC) does not exceed 200,000 Bitcoins (see Fig. 2), and volume does not change significantly as compared to the price change and Google Trends Index change.

5.4. A binomial measure from P-model

Based on Podobnik and Stanley [13], when $q = 2$, for two fractionally autoregressive integral moving average (ARFIMA) processes that share the same random noise, the cross-correlation exponent is approximately identical to the mean value of the individual Hurst exponents. Zhou [16] also observed that for two time series constructed by a binomial measure from

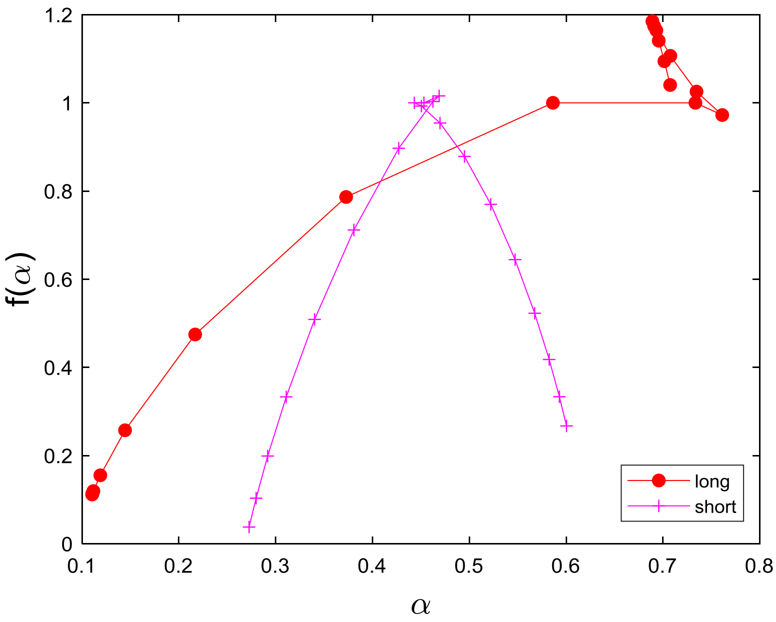

Fig. 13. The multifractal spectrum for *CGT* and *Ret* both in long-term and short-term.

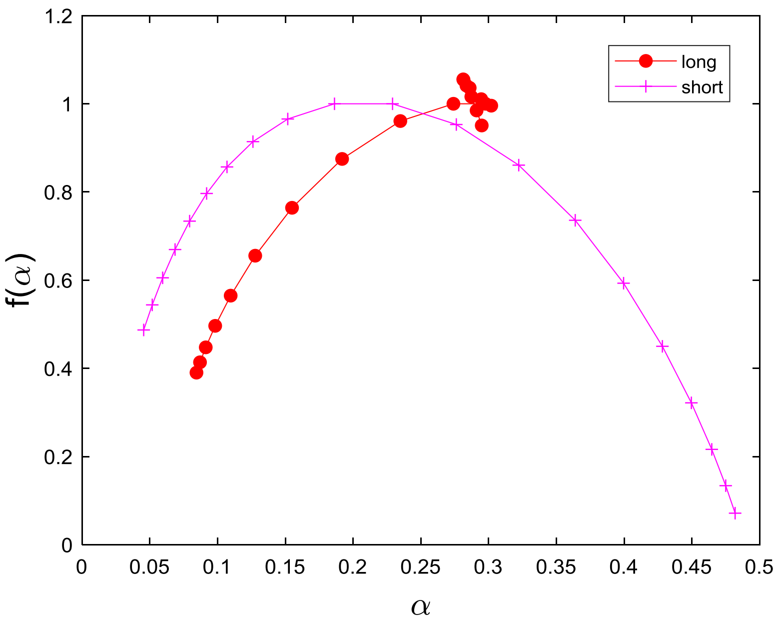

Fig. 14. The multifractal spectrum for *CGT* and *CV* both in long-term and short-term.

the *p*-model, the following equation exists

$$H_{xy}(q) = \left[H_{xx}(q) + H_{yy}(q)\right]/2 \tag{18}$$

The general Hurst exponents $H_{xx}(q)$ and $H_{yy}(q)$ for cross-correlation between *CGT* and *Ret* as well as *CGT* and *CV* are calculated by using MF-DFA. If $H_{xx}(q)$ and $H_{yy}(q)$ do not depend on the change of q, $H_{xx}(q)$ and $H_{yy}(q)$ do not have the characteristics of multifractal. As we can see in Fig. 16, Google Trends and Bitcoin have multifractal features. Hence, the average scaling exponents are the black lines with asterisk in Fig. 16. For the cross-correlation between *CGT* and *CV*, we find that the mean scaling exponents are larger than the cross-correlation exponents for q <-1 and smaller when q> -1. For cross-correlation between *CGT* and *Ret*, the mean scaling exponents are very close to the cross-correlation exponents.

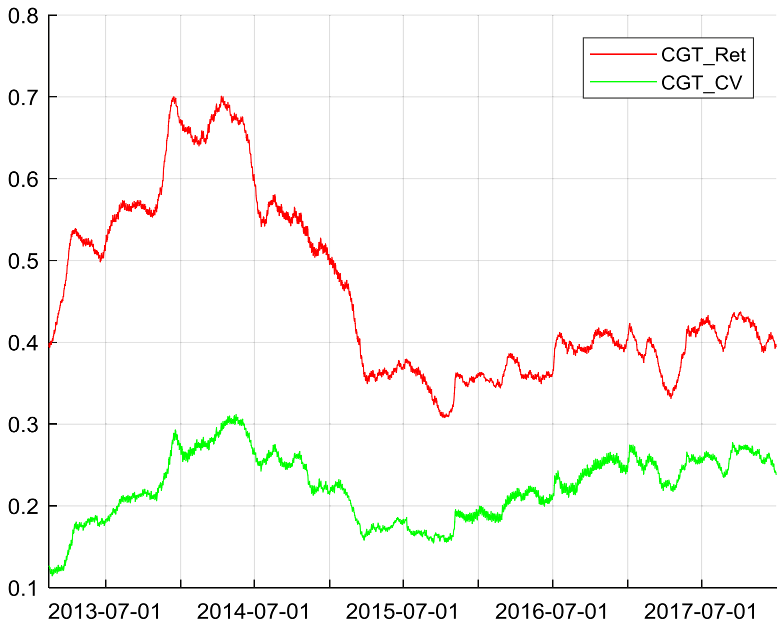

Fig. 15. Time varying scaling exponents when $q = 2$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

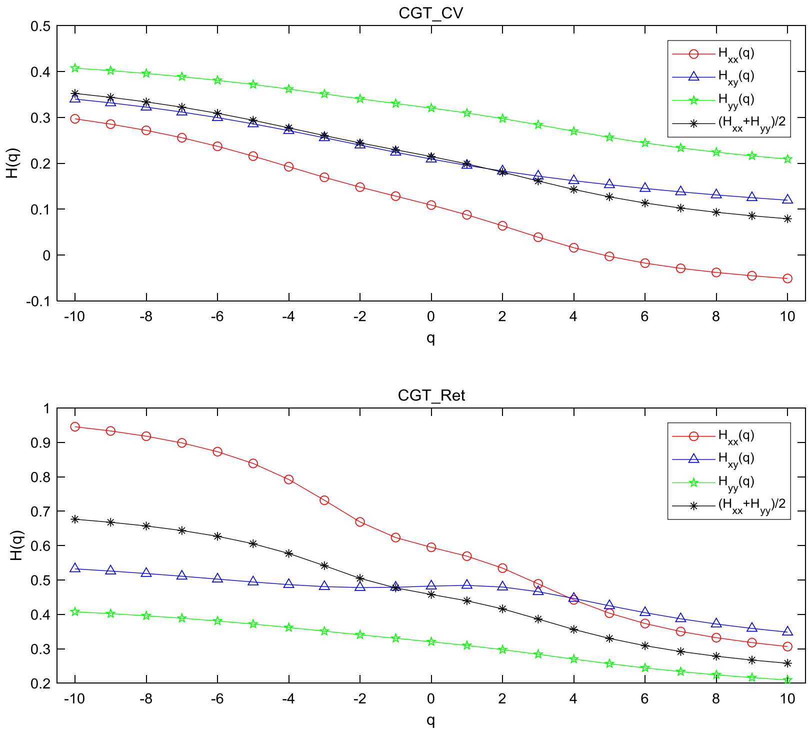

Fig. 16. The relationship of $H_{xx}(q)$, $H_{xy}(q)$, $H_{yy}(q)$, $\left[H_{xx}(q) + H_{yy}(q)\right]/2$ and different q.

6. Conclusion

In this paper, we quantify the cross-correlations between change of Google Trends and Bitcoin (returns and change of volume) with MF-DCCA method. The empirical results mainly show that: First, based on the significance of statistical analysis, we find that cross-correlation is generally significant by using Q_{cc} test. Second, there are power-law cross-correlations between change of Google Trends and Bitcoin (returns and change of volume). The cross-correlations all perform multifractality. We find that the cross-correlation between *CGT* and *Ret* has a strong multifractal in the long-term, but the short-term multifractal is weak. The cross-correlation between *CGT* and *CV* is the opposite. Moreover, the cross-correlations between *CGT* and *Ret* of small fluctuations are persistent in the long-term and those of large fluctuations are anti-persistent in the short-term while the cross-correlations between *CGT* and *CV* of large and small fluctuations are anti-persistent both in the short-term and long-term. Third, we find that there is a decrease trend of cross correlation scaling exponents. In the entire period, the cross-correlations between *CGT* and *CV* is anti-persistent, while the cross-correlations between *CGT* and *Ret* is persistent cross-correlated when Bitcoin had just begun not long time.

Admittedly, the above results provide conclusions from a multifractal perspective and do not fully disclose the profound relationship between emerging assets and online activities. Since Bitcoin does not have the fundamental determinants of stocks and foreign exchange, behavior of Bitcoin's price and volume has complex dynamics. For a more in-depth study, we also need to consider Bitcoin forum discussions, policy restrictions, changes in the blockchain used by Bitcoin (hard-forking), and the geographic location of the trading platform. We leave these for future research.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (71701150 and 71790594) and Young Elite Scientists Sponsorship Program by Tianjin (TJSQNTJ-2017-09) and Fundamental Research Funds for the Central Universities (63182064).

References

[1] G.P. Smith, Google Internet search activity and volatility prediction in the market for foreign currency, Finan. Res. Lett. 9 (2) (2012) 103–110.
[2] L. Bijl, G. Kringhaug, P. Molnár, E. Sandvik, Google searches and stock returns, Int. Rev. Finan. Anal. 45 (2016) 150–156.
[3] D. Shen, Y. Zhang, X. Xiong, W. Zhang, Baidu index and predictability of Chinese stock returns, Financ. Innov. 3 (1) (2017) 4.
[4] T. Preis, H.S. Moat, H.E. Stanley, Quantifying trading behavior in financial markets using Google Trends, Sci. Rep. 3 (2013) srep01684.
[5] L. Kristoufek, What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis, PLoS One 10 (4) (2015) e0123923.
[6] P. Ciaian, M. Rajcaniova, Kancs d.A., The economics of BitCoin price formation, Appl. Econ. 48 (19) (2016) 1799–1815.
[7] A. Urquhart, The inefficiency of Bitcoin, Econ. Lett. 148 (2016) 80–82.
[8] S. Nadarajah, J. Chu, On the inefficiency of Bitcoin, Econ. Lett. 150 (2017) 6–9.
[9] L. Kristoufek, On Bitcoin markets (in) efficiency and its evolution, Physica A (2018).
[10] A.F. Bariviera, M.J. Basgall, W. Hasperué, M. Naiouf, Some stylized facts of the Bitcoin market, Physica A 484 (Supplement C) (2017) 82–90.
[11] D. Garcia, C.J. Tessone, P. Mavrodiev, N. Perony, The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy, J. R. Soc. Interface 11 (99) (2014) 20140623.
[12] D.G. Baur, T. Dimpfl, K. Kuck, Bitcoin, gold and the US dollar – A replication and extension, Finan. Res. Lett (2018).
[13] B. Podobnik, H.E. Stanley, Detrended cross-correlation analysis: A new method for analyzing two nonstationary time series, Phys. Rev. Lett. 100 (8) (2008) 084102.
[14] B. Podobnik, Z.-Q. Jiang, W.-X. Zhou, H.E. Stanley, Statistical tests for power-law cross-correlated processes, Phys. Rev. E 84 (6) (2011) 066118.
[15] J.W. Kantelhardt, S.A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, H.E. Stanley, Multifractal detrended fluctuation analysis of nonstationary time series, Physica A 316 (1) (2002) 87–114.
[16] W.-X. Zhou, Multifractal detrended cross-correlation analysis for two nonstationary signals, Phys. Rev. E 77 (6) (2008) 066211.
[17] A. Siganos, E. Vagenas-Nanos, P. Verwijmeren, Facebook's daily sentiment and international stock markets, J. Econ. Behav. Organ. 107 (2014) 730–743.
[18] W. Zhang, X. Li, D. Shen, A. Teglio, Daily happiness and stock returns: Some international evidence, Physica A 460 (2016) 201–209.
[19] X. Jin, D. Shen, W. Zhang, Has microblogging changed stock market behavior? Evidence from China, Physica A 452 (2016) 151–156.
[20] X. Li, D. Shen, W. Zhang, Do Chinese internet stock message boards convey firm-specific information? Pacific-Basin Finance J. 49 (2018) 1–14.
[21] Z. Da, J. Engelberg, P. Gao, The sum of all FEARS investor sentiment and asset prices, Rev. Financ. Stud. 28 (1) (2014) 1–32.
[22] L. Kristoufek, Can Google Trends search queries contribute to risk diversification? Sci. Rep. 3 (2013) 2713.
[23] D. Shen, X. Li, W. Zhang, Baidu news information flow and return volatility: Evidence for the sequential information arrival hypothesis, Econ. Modell (2017).
[24] Open source information, investor attention, and asset pricing, Econ. Modell. 33 (2013) 613–619.
[25] T. Dimpfl, S. Jank, Can internet search queries help to predict stock market volatility? Europ. Finan. Manage. 22 (2) (2016) 171–192.
[26] Z. Da, J. Engelberg, P. Gao, In search of attention, J. Finance 66 (5) (2011) 1461–1499.
[27] N. Vozlyublennaia, Investor attention, index performance, and return predictability, J. Bank. Financ. 41 (2014) 17–35.
[28] M.S. Drake, D.T. Roulstone, J.R. Thornock, Investor information demand: Evidence from Google searches around earnings announcements, J. Acc. Res. 50 (4) (2012) 1001–1040.
[29] A.F. Bariviera, The inefficiency of Bitcoin revisited: a dynamic approach, Econ. Lett. 161 (2017) 1–4.
[30] J. Alvarez-Ramirez, E. Rodriguez, C. Ibarra-Valdez, Long-range correlations and asymmetry in the Bitcoin market, Physica A 492 (Supplement C) (2018) 948–955.
[31] D.G. Baur, K. Hong, A.D. Lee, Bitcoin: Medium of exchange or speculative assets? J. Int. Finan. Markets, Inst. Money (2018) Forthcoming.
[32] J.R. Hendrickson, W.J. Luther, Banning bitcoin, J. Econ. Behav. Organ. 141 (Supplement C) (2017) 188–195.
[33] T. Kim, On the transaction cost of Bitcoin, Finan. Res. Lett. 23 (Supplement C) (2017) 300–305.
[34] N. Gandal, J.T. Hamrick, T. Moore, T. Oberman, Price manipulation in the Bitcoin ecosystem, J. Monet. Econ. (2018) Forthcoming.
[35] L. Kristoufek, BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era, Sci. Rep. 3 (2013) 3415.
[36] A. Yelowitz, M. Wilson, Characteristics of Bitcoin users: an analysis of Google search data, Appl. Econ. Lett. 22 (13) (2015) 1030–1036.
[37] E. Alessio, A. Carbone, G. Castelli, V. Frappietro, Second-order moving average and scaling of stochastic time series, Eur. Phys. J. B 27 (2) (2002) 197–200.
[38] Z.-Q. Jiang, W.-J. Xie, W.-X. Zhou, Testing the weak-form efficiency of the WTI crude oil futures market, Physica A 405 (Supplement C) (2014) 235–244.
[39] L. Kristoufek, Detrending moving-average cross-correlation coefficient: Measuring cross-correlations between non-stationary series, Physica A 406 (2014) 169–175.

[40] B. Podobnik, D. Horvatic, A.M. Petersen, H.E. Stanley, Cross-correlations between volume change and price change, Proc. Natl. Acad. Sci. 106 (52) (2009) 22079–22084.
[41] L. Kristoufek, Measuring correlations between non-stationary series with DCCA coefficient, Physica A 402 (2014) 291–298.
[42] L. Kristoufek, Power-law correlations in finance-related Google searches, and their cross-correlations with volatility and traded volume: Evidence from the Dow Jones Industrial components, Physica A 428 (2015) 194–205.
[43] G.-F. Gu, W.-X. Zhou, Detrending moving average algorithm for multifractals, Phys. Rev. E 82 (1) (2010) 011136.
[44] H.-C. Xu, G.-F. Gu, W.-X. Zhou, Direct determination approach for the multifractal detrending moving average analysis, Phys. Rev. E 96 (5) (2017) 052201.
[45] Z.-Q. Jiang, W.-X. Zhou, Multifractal detrending moving-average cross-correlation analysis, Phys. Rev. E 84 (1) (2011) 016106.
[46] L. Kristoufek, Multifractal height cross-correlation analysis: A new method for analyzing long-range cross-correlations, Europhys. Lett. 95 (6) (2011) 68001.
[47] J. Wang, P. Shang, W. Ge, Multifractal cross-correlation analysis based on statistical moments, Fractals 20 (03n04) (2012) 271–279.
[48] C. Meneveau, K. Sreenivasan, P. Kailasnath, M. Fan, Joint multifractal measures: Theory and applications to turbulence, Phys. Rev. A 41 (2) (1990) 894.
[49] W.-J. Xie, Z.-Q. Jiang, G.-F. Gu, X. Xiong, W.-X. Zhou, Joint multifractal analysis based on the partition function approach: analytical analysis, numerical simulation and empirical application, New J. Phys. 17 (10) (2015) 103020.
[50] P. Oświęcimka, S. Drożdż, M. Forczek, S. Jadach, J. Kwapień, Detrended cross-correlation analysis consistently extended to multifractality, Phys. Rev. E 89 (2) (2014) 023305.
[51] J. Kwapień, P. Oświęcimka, S. Drożdż, Detrended fluctuation analysis made flexible to detect range of cross-correlated fluctuations, Phys. Rev. E 92 (5) (2015) 052815.
[52] X.-Y. Qian, Y.-M. Liu, Z.-Q. Jiang, B. Podobnik, W.-X. Zhou, H.E. Stanley, Detrended partial cross-correlation analysis of two nonstationary time series influenced by common external forces, Phys. Rev. E 91 (6) (2015) 062816.
[53] Z.-Q. Jiang, X.-L. Gao, W.-X. Zhou, H.E. Stanley, Multifractal cross wavelet analysis, Fractals 25 (06) (2017) 1750054.
[54] Z.-Q. Jiang, Y.-H. Yang, G.-J. Wang, W.-X. Zhou, Joint multifractal analysis based on wavelet leaders, Front. Phys. 12 (6) (2017) 128907.
[55] P. Yue, H.-C. Xu, W. Chen, X. Xiong, W.-X. Zhou, Linear and nonlinear correlations in the order aggressiveness of Chinese stocks, Fractals 25 (05) (2017) 1750041.
[56] F. Ma, Q. Zhang, C. Peng, Y. Wei, Multifractal detrended cross-correlation analysis of the oil-dependent economies: Evidence from the West Texas intermediate crude oil and the GCC stock markets, Physica A 410 (2014) 154–166.
[57] X. Zhuang, Y. Wei, B. Zhang, Multifractal detrended cross-correlation analysis of carbon and crude oil markets, Physica A 399 (2014) 113–125.
[58] L. Kristoufek, Testing power-law cross-correlations: Rescaled covariance test, Eur. Phys. J. B 86 (10) (2013) 418.
[59] B. Podobnik, I. Grosse, D. Horvatić, S. Ilic, P.C. Ivanov, H.E. Stanley, Quantifying cross-correlations using local and global detrending approaches, Eur. Phys. J. B 71 (2) (2009) 243–250.
[60] R. Balocchi, M. Varanini, A. Macerata, Quantifying different degrees of coupling in detrended cross-correlation analysis, Europhys. Lett. 101 (2) (2013) 20011.
[61] D.A. Blythe, V.V. Nikulin, K.-R. Müller, Robust statistical detection of power-law cross-correlation, Sci. Rep. 6 (2016) 27089.
[62] G.M. Ljung, G.E. Box, On a measure of lack of fit in time series models, Biometrika 65 (2) (1978) 297–303.
[63] F. Ma, Y. Wei, D. Huang, Multifractal detrended cross-correlation analysis between the Chinese stock market and surrounding stock markets, Physica A 392 (7) (2013) 1659–1670.
[64] Y. Wang, Y. Wei, C. Wu, Cross-correlations between Chinese A-share and B-share markets, Physica A 389 (23) (2010) 5468–5478.