MScF Degree Thesis

AY 2023

TESTING FOR NONLINEAR PROPERTIES AND CHAOS PHENOMENON OF BITCOIN

57211527 LIN, YU CHI

HOT PAPERS IN FINANCE

C.E. PROF. COAD, ALEX

D.E. PROF. ARIKAWA, YASUHIRO D.E. PROF. CHEUNG, MING YAN WILLIAM

Summary

Since the 1970s, the Efficient Market Hypothesis has dominated financial theory, backed by a plethora of empirical research. Yet, the prediction of asset prices continues to be a central focus for academics. The introduction of the theory of implicit nonlinear processes in the market has led to empirical outcomes that often contradict previous findings. As a result, the quest to identify patterns within the stochastic behaviors of short-term and long-term price fluctuations has emerged as a much debated and focused area of investigation.

This study primarily employs the "BDS statistic," a concept introduced by Brock, Dechert, and Scheinkman in 1987, to investigate the dynamics and traits of the daily closing data of the hourly Bitcoin return series and volatility series from January 2021 to December 2021. The goal is to determine whether these dynamics align with the phenomenon of nonlinear deterministic chaos. Considering the widespread use of the ARMA-GARCH models in the field of finance, this study considers the fitting of an ARMA-GARCH model to the return series of Bitcoin, and delves into whether said model can adequately account for the fluctuations and dynamics of Bitcoin.

The findings of the study suggest that when the ARMA-GARCH model is fitted to the Bitcoin return series, multiple test results indicate a good fit that accounts for autocorrelation and heteroskedasticity. Nevertheless, the inaptitude at pinpointing a concrete pattern that elucidates the volatility behavior of the time series at hand and the irregularity of the distribution of the dataset may have contributed to the inability to fully capture all the characteristics that determine the price changes of Bitcoin, indicating the inevitability of the model carrying inherent forecasting risk. From the study, we also see that the Bitcoin market does not reflect market efficiency, and the Bitcoin volatility series adheres to a nonlinear stochastic process as opposed to a nonlinear deterministic process. This indicates that there may be chaotic components that govern the underlying system. These findings also show that the price changes of Bitcoin have nonlinear characteristics, while also having clustering and long-term memory properties.

1. Introduction:

Section 1.1: Research Motivation

Since the 1970s, the efficient market hypothesis has become the mainstream of financial theory, supported by numerous empirical studies (Fama, 1970). If the efficient market hypothesis holds true, then all models predicting price movements would become ineffective. This is because no model can quickly respond to changes in all external variables and achieve efficient predictions. However, technical analysis opposing this hypothesis still exists, indicating the presence of phenomena that the current hypothesis cannot explain.

With the introduction of new market hypothesis, namely the market's implied nonlinear process, the empirical results obtained have diverged significantly from previous findings (Scheinkman & LeBaron, 1989; Larrain, 1991; LeBaron, 1991). In these related studies, nonlinear dynamics have been found in various financial and economic domains, such as the labor market and investment environment in the United States, as highlighted by Brock and Sayers (1988). Therefore, attempting to identify the possible regularities embedded in the randomness of market prices in the short and long terms has emerged as a much debated and focused area of investigation. The idea of testing out this hypothesis has been a recurring theme with each trending financial asset on the market, be it the index funds of the 1980s, the exchange-traded funds of the 1990s, or the structured products, such as collateralized debt obligations of the early 2000s.

As of the beginning of the 2010s, cryptocurrencies have surfaced as a conspicuous and electrifying asset class, commanding substantial scrutiny in the vast expanse of global finance. With decentralized digital currencies like Bitcoin and Ethereum skyrocketing in popularity, both governments and financial institutions alike have begun to recognize their significance. The decentralized essence and potential for substantial gains have enraptured the attention of investors, enthusiasts, and speculators alike. The assessment of cryptocurrencies has grown increasingly pivotal, given their inherent capriciousness and the myriad of applications they proffer. Scrutinizing market trends, regulatory breakthroughs, and technological strides assume a paramount role in apprehending the capricious intricacies of the cryptocurrency ecosystem.

Out of the many cryptocurrencies available on the crypto market, Bitcoin, an innovative digital currency, has been at the very forefront of the revolution of the financial field ever since its inception in 2009. Launched by the enigmatic individual or collective known as Satoshi Nakamoto, Bitcoin has the notion of decentralized, peer-to-peer transactions built upon blockchain technology. Its multifaceted attributes, momentous historical journey, and profound impact within the financial domain have fundamentally transformed our perception and interaction with monetary systems.

Bitcoin operates as a decentralized cryptocurrency, emancipated from the clutches of central authorities such as banks or governments. It thrives on the blockchain, which meticulously documents all transactions atop it, ensuring unparalleled transparency, impregnability, and immutability. The verification of these transactions falls upon a group of network participants, known as miners, who harness computational power to decipher and append fresh blocks to the blockchain. In recognition of their endeavors, miners are rewarded with freshly minted bitcoins.

The core essence of Bitcoin lies in its unparalleled capacity to transcend the inherent boundaries of traditional financial structures. By obliterating the more traditional intermediaries, such as banks, Bitcoin empowers direct peer-to-peer transactions, greatly reducing transaction fees and processing durations. Furthermore, it affords a modicum of pseudonymity, employing cryptographic addresses rather than personal information to identify users. This particular facet has attracted privacy-seeking individuals and those ensnared within countries harboring restrictive financial systems.

The history of Bitcoin is punctuated by momentous milestones and pivotal events. Initially, it garnered traction predominantly among tech-savvy aficionados. Over time, Bitcoin's popularity skyrocketed, engendering volatility in its value and spurring speculative trading. In 2017, its worth reached unprecedented heights, with a solitary bitcoin surpassing the $19,000 mark before experiencing subsequent corrections. The impact of Bitcoin upon the financial field is momentous and far-reaching. As a decentralized currency, it perpetually challenges the orthodox banking system by presenting an alternative store of value and medium of exchange. The finite supply of 21 million bitcoins has led some to perceive it as a digital equivalent to gold. Consequently, individuals and institutions have pondered allocating a fraction of their investment portfolios to Bitcoin, viewing it as a bulwark against inflation or economic instability.

Bitcoin has fostered financial inclusion by extending banking services to the unbanked or underbanked populations. Through the utilization of smartphones, individuals can actively participate in the global economy and partake in cross-border transactions while foregoing the need for a conventional bank account. This development holds particular significance for denizens of developing nations, where accessibility to financial services remains constricted. The advent of Bitcoin has also sparked a surge of innovation in the form of blockchain technology. The underlying blockchain infrastructure has transcended the confines of cryptocurrencies, enticing various industries to explore its boundless potential. Blockchain promises secure, transparent, and efficient record-keeping, supply chain management, smart contracts, and beyond. Its tendrils extend far beyond the realm of finance, permeating sectors such as healthcare, logistics, and governance.

Bitcoin confronts a variety of challenges. Its price volatility has instigated concerns pertaining to market manipulation and the latent threat of financial instability. Regulatory frameworks for cryptocurrencies remain in a state of flux, as governments worldwide grapple with the arduous task of classifying and regulating Bitcoin and other digital assets. The energy consumption associated with Bitcoin mining has also garnered criticism due to its environmental impact, prompting deliberations on the necessity for sustainable alternatives. Bitcoin has emerged as a transformative force within the financial arena, introducing the revolutionary concept of decentralized currency and blockchain technology. Its attributes of decentralization, transparency, and pseudonymity have posed formidable challenges to conventional financial systems, igniting discourse on the future trajectory of monetary systems. The significance, historical voyage, and pervasive impact of Bitcoin extend has grown and rooted itself as one of the major players of the financial sector.

Section 1.2: Research Objectives

Initially, the introduction of cryptocurrencies was aimed at creating decentralized and transparent financial systems. However, as reported in the working paper by the Bank of International settlements in November 2022, staggering “estimations that 73-81% of global investors have likely lost money on their crypto investment” (Raphael, 2022). In addition, in the past couple of years, global regulatory bodies have stringently kept the cryptocurrency market capitalization under control, predicting the potential cash flow to be directed into the speculative trading of cryptocurrencies in the future. These attempted interferences in the forms of regulations and monitors within the financial field have caused ripples and shocks throughout the cryptocurrency market. Amidst these recent international developments and shifts in the global financial field, this study sets out to examine the characteristics of the hourly closing data of Bitcoin. The purpose of this paper can be categorized into brief summaries of the research objectives which are discussed in the following paragraphs.

The first objective is to understand the properties Bitcoin time series through analyzing the trends, seasonality, and volatility clustering. Tests on whether the fluctuations of the hourly Bitcoin closing price series conforms to the efficient market hypothesis would also be conducted. If the behavior of the hourly Bitcoin closing series were to conform to the efficient market hypothesis, as mentioned before in the research motivation in Section 1, any model that aims to predict price fluctuations of Bitcoin would be rendered ineffective. If the behavior does not conform to the efficient market hypothesis, it would indicate that the past bitcoin prices hold certain components and factors that affect the future price of Bitcoin and its fluctuation, which can assist in formulating a prediction model. In this were the case, we will further study the characteristics of the Bitcoin price fluctuations.

The second objective is to verify the possible existence of potential non-linear processes in the Bitcoin time series. This would be conduct by fitting an appropriate ARMA-GARCH model to first capture and eliminate the linear dynamics and volatility patterns in the Bitcoin time series. Also, various tests assessing the adequacy of the fit would be conducted. When attempting to dissect a time series for its characteristics, it is necessary to distinguish whether the relationship is linear or non-linear. This will aid in the fitting of the model and help us better understand the underlying processes, reduce the possible chance of overfitting or underfitting, and choosing the most effective tool for analyzing the time series. This study would like to analyze if an ARMA-GARCH model has effectively captured all the dynamics and accounted for the fluctuations of the time series.

Last, we test for the possibility of chaos phenomenon existing within the residuals and the volatility of the hourly Bitcoin return series. The chaos phenomenon can be captured with a meticulous series of tests, accounting for whether the series has long-term dependency and comparing their diffusion rate to a geometric Brownian motion. If the mentioned time series indeed exhibit properties of chaos, an attempt will be made to identify the number of dynamical correlated variables for this fluctuation pattern in order to establish an idea of what the underlying system may look like. If no chaotic phenomenon can be picked up, then it would indicate that the fluctuating behavior of the time series cannot be effectively predicted by a single model or system.

Section 1.3: Data

Before looking at the data chosen, the reasons and significance as to why this particular time frame is selected will be explored. Bitcoin, by nature, is a rather young and highly speculative asset. Its volatility is driven by various factors including, but not limited to, regulatory news, market sentiment, technological advancements, and macroeconomic trends, quite similar with other high volatility financial assets on the market. These diverse factors introduce complex, nonlinear dynamics into the Bitcoin market, and to a large extent, the overall cryptocurrency market. Studying these can highlight the specific conditions or factors that trigger major price swings. The enigmatic undulations in Bitcoin's price allude to the possibility that certain pivotal catalysts do not adhere to a rudimentary, linear correlation with its value. This enigmatic behavior intimates that the Bitcoin market responds to specific stimuli in an exponential or other intricately convoluted mode. The intricate dynamics of Bitcoin's price movements imply that the underlying forces at play are far from straightforward and call for a deeper understanding of the market.

The hourly data is specifically chosen to capture the intraday volatility of Bitcoin. Given the global nature of the Bitcoin market, which operates 24/7, news events or market sentiments can shift drastically within a day, leading to rapid changes in prices. Opting for an hourly time series presents a vivid and dynamic portrayal of Bitcoin's price fluctuations. By leveraging hourly data, the abundance of data points gained allows for a more intricate and nuanced understanding of the volatility in Bitcoin prices compared to daily, weekly, or monthly data, potentially providing more data points and thus resulting in a richer and more detailed perspective on the fluctuations in Bitcoin prices. The utilization of shorter timescales enables the capture of swift shifts in investor sentiment, abrupt market responses to news events, and other intraday dynamics which can be better captured and analyzed using hourly data while eluding detection in longer timescales. Shorter timescales can capture rapid changes in investor sentiment, sudden market reactions to news events, and other quick fluctuations that may get smoothed out or unnoticed in longer timescales.

As stated before, a core concept of chaos theory lies in finding order in seemingly random data. Financial markets frequently exhibit non-linear properties, signifying that the outcome is not simply proportional to the input. Likewise, we can compare these chaotic behaviors, which can be seen as a more complex expression of nonlinearity, of the traditional financial market to that of the Bitcoin and cryptocurrency market, further suspecting a similar process to be at play, where minute alterations in the initial conditions yield astounding and unforeseeable results. When it comes to Bitcoin, comprehending these characteristics becomes pivotal for constructing more precise and robust prediction models. Financial time series, akin to the fluctuating movements of Bitcoin prices, frequently reveal themselves as non-linear and chaotic. Being able to grasp these signals would be the most importance piece in crafting more refined forecasting models. Comprehending the dynamics of such turbulent systems can also uncover the underlying processes that lead to the changes in the asset's value.

As for the year component, 2021 is a particular interesting year for Bitcoin considering all that has happened. The mercurial price of Bitcoin in 2021 exhibited extraordinary fluctuations, with numerous momentous surges and plunges. Delving into this year's data could grant unparalleled revelations about the intricate interplay between specific occurrences and the broader economic events, and their profound influence and implications on the market. The year 2021 was teeming with impactful events that left an indelible imprint on Bitcoin's price, ranging from Tesla's entanglement to China's crackdown on crypto mining. Scrutinizing the chaos and non-linear dynamics in Bitcoin's price during this period could yield invaluable insights into how these external events and factors metamorphose into dramatic price gyrations within the cryptocurrency market. The year started off with low closing prices of around 30000 in January, which can be in part be attributed to the effects of the raging pandemic, then skyrocketing to an all-time high in April, to nearly 65000. Bitcoin will later in late May fall all the way back to 35000, and then making a comeback in October, successfully breaking through the 65000 mark. These large fluctuations in price makes 2021 a prime candidate for studying the more sensitive characteristics of Bitcoin to shocks in the environment and how the price of Bitcoin is affected by these shocks. The data selected for the analysis of this study are the hourly closing prices of Bitcoin from January 1, 00:00:00, 2021 to January 1, 00:00:00, 2022, coming to a total of 8760 observations. These observations were taken from the database of Bitstamp, the unit in use for the pricing of Bitcoin is the US dollar.

2. Prior Research:

This chapter consists of two sections. In the first section, we provide an overview of chaos theory, the implications, and sum up the properties of it from prior research. Since the number of studies that include the applications of chaos theory to cryptocurrency remain relatively low, the second section would focus mainly on reviewing and discussing empirical studies related to chaos theory in stock markets, through which we may get a sense of how the techniques can be applied to the hourly Bitcoin time series.

Section 2.1: Chaos Theory

Chaos describes a seemingly random non-linear deterministic process that captures the behavior of a variable over a period of time. Despite its name and its phenomenon seeming to be random, it is actually a deterministic phenomenon, meaning that if all the underlying pattern is known, the behavior of the variable in question can be fully explained. Chaos theory got its name because it studies these seemingly random or chaotic situations which are, in fact, governed by underlying patterns and deterministic laws. To the casual observer, the behavior may appear disorderly and unpredictable, but it can be mathematically described by deterministic equations. In contrast, values generated by a random variable cannot be accurately predicted. Some crucial aspects of chaos theory include high sensitivity to initial conditions, inherent deterministic nature, fractal geometrical, and attractors.

High sensitivity to initial conditions is probably the most recognized aspect of chaos theory, also frequently known as the butterfly effect. This concept suggests that a small change, such as a butterfly flapping its wings in Brazil, can ultimately set off a chain of events leading to a significant occurrence, like a hurricane on the other side of the planet. Such implications when applied to the fields of mathematics or physics systems would indicate that minor changes in initial conditions can lead to significantly divergent results, rendering long-term prediction impracticable in general terms.

The deterministic nature of chaos phenomenon is presented in the sense that despite their seemingly random behavior, they follow deterministic rules that are clearly defined, as stated earlier. In other words, the future states of the system at hand, given a certain state, are entirely determined by the system's governing laws. The seemingly randomness of the process and outcome is not a product of randomness within the system's governing laws, but of the system's high sensitivity to initial conditions.

As for the fractal geometrical properties of the chaos phenomenon would best be understood in mathematics, where a fractal is a geometric construct that displays self-similarity, meaning it maintains the same appearance regardless of scale. If one were to zoom into a fractal, he would observe the repetition of similar patterns at every level of magnification. Many chaotic systems display fractal behavior where the system's pattern recurs consistently across various scales.

The attractors of chaos phenomenon refer to a set of numerical values within a dynamic system to which a system tends to evolve, irrespective of its initial state. Chaotic systems often exhibit what are known as strange attractors, distinct from simple point and periodic attractors. A strange attractor possesses a fractal nature, and its fractal dimension does not align with an integer value. The above-mentioned characteristics of chaos theory have attracted many a scholar from various fields who are interested in applying the theory to their respective domains of expertise.

Within literatures on business cycles, two scenarios are mentioned that may cause variations in future values. One is the Box-Jenkins time series model, which assumes stable but unbalanced economic conditions. When the economy is subject to external shocks such as wars or natural disasters, they are reflected in future values. In this model, certain dynamic economic behaviors are a response to these external shocks. The other scenario is chaos growth theory, a non-linear dynamic economic theory that suggests economic fluctuations originate from within the system itself. Since chaotic dynamics are inherently non-linear processes and exhibit more complex patterns compared to linear models, researchers turned to this area of study after the significant price volatility in the U.S. stock market on October 19, 1987, to seek explanations for large-scale asset price movements (Hsieh, 1991).

In this study, we illustrate the chaos processes using the logistic growth equation, which is a representative equation. The equation is as follows:

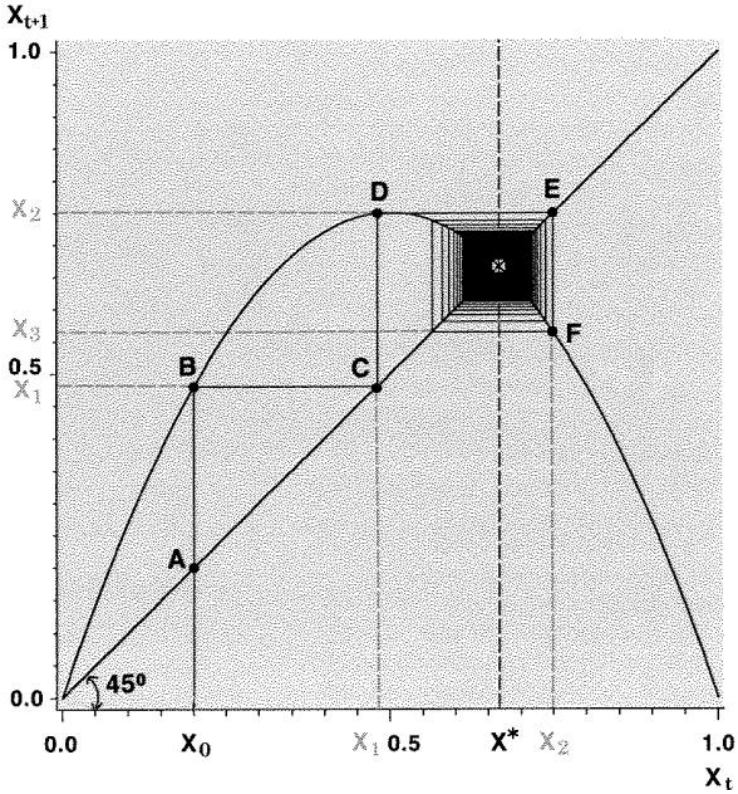
$$X_{t+1} = kX_t (1 – X_t)$$

$$0 < X_t <1 , 0 < k < 4$$

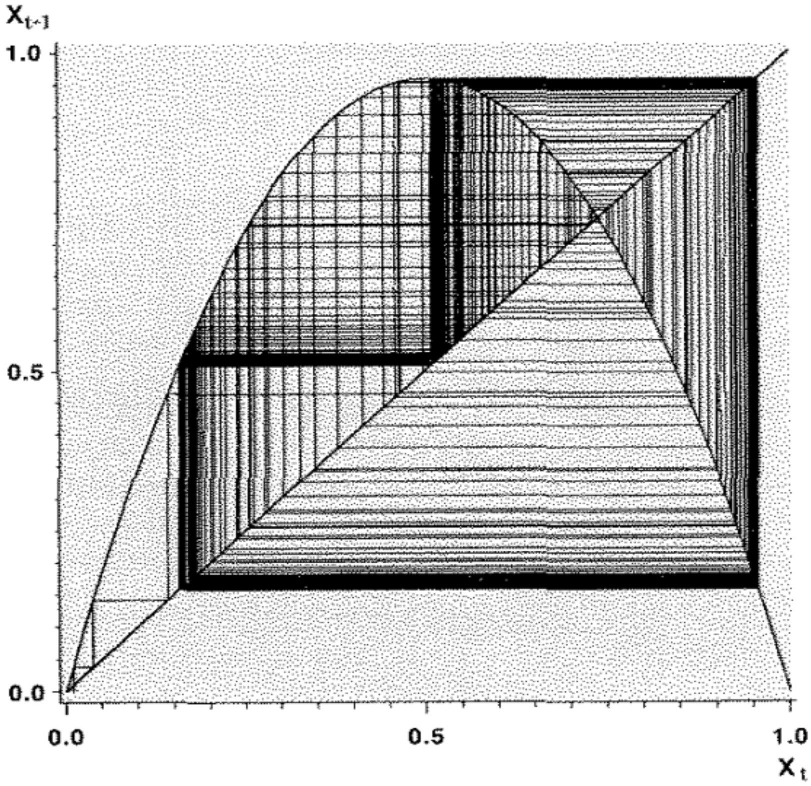
By exploring this equation, we investigate and summarize the properties of chaos. The equation as shown above describes the path of the variable X over time, where X_{t+1} is a function that includes the previous value X_t and the parameter k. The value of X is a real number between 0 and 1. The steepness of the parabola is determined by the sole parameter k, referred to as the "tuning parameter". A larger value of k leads to a greater absolute rate of change in X. When the tuning parameter k is less than or equal to 3, regardless of the initial value, the system reaches a stable equilibrium.

To illustrate this, consider the case where k = 3 and the initial value X_0 = 0.20. After 500 time units, the logistic growth curve converges to a stable equilibrium point (see Figure 2-1). In the figure, the parabola represents the quadratic equation $X_{t+1} = 3X_t (1 – X_t)$, and all values of X_t and X_{t+1} fall on this curve. The 45-degree line represents the set of points where $X_{t+1} = X_t$, which corresponds to the condition of reaching a stable equilibrium. Point B in the figure represents the

$$X_{t+1} = 3\, X_t (1 - X_t)$$
$$X_0 = .20 \quad t = 1 \text{ to } 500$$



$$X_{t+1} = 3.82840\, X_t (1 - X_t)$$
$$X_0 = .0101 \quad t = 1 \text{ to } 500$$



initial value $X_0 = 0.20$, which reacts to the function and yields a response value of 0.48. Point C represents the transformation of the response value of X_0 to the input value of X_1. By applying the function again, we obtain another response value of 0.7488, represented by point D. By repeating the above process, we obtain points E and F. Through continuous iterations, the equation eventually converges to a point in the figure, which has a value of 2/3 and serves as the stable solution for this equation.

When k exceeds 3, the solution of the function becomes an unstable equilibrium point. The variable oscillates between two points, creating a cyclic pattern. As k increases, the system generates more unstable equilibrium points. The trajectory of the function exhibits a cycle of four points, then eight points, and continues to increase until it forms a cycle of 2^n points where (n = 1, 2, 3, ...). This process of increasing cycles is referred to as bifurcation (May, 1976).

When the value of k reaches 3.57, the system enters the range where chaos may occur ($3.57 < k < 4$). Within this range where there may be chaos, the function exhibits two types of trajectories. One is a cyclic oscillation with an infinite number of points, and the other is a path without periodicity.

Figure 2-2 presents an example of non-periodic chaos. No matter how many times this equation is executed, it is impossible to generate the same value of X_t. If these output values are plotted in a time series graph, they may be mistaken for values generated by a random process. However, this pattern does not include any random components; its coefficients are entirely determinable through the underlying equation. Feigenbaum (1983) demonstrated that the value of k is independent of the curve's shape. As long as the function curve y = f(x) has a local maximum, when k approximates 3.570, the curve's trajectory exhibits chaotic behavior. In other words, the value of k serves as a universal constant in nature.

Summarizing what we get from the above, the properties of chaos can be summarized as the four following points:

1. Chaos is a non-linear process.
2. Chaos is highly sensitive to initial values and tuning parameters.
3. Chaos cannot produce identical output values and may have numerous stable equilibrium points.
4. The correlation dimension of chaos is not constant (Grassberger & Procaccia, 1983).

Section 2.2: Applications of chaos theory on stock markets

In this section, the main focus is to investigate the empirical results obtained from the application of various testing tools related to chaos theory in stock markets and cryptocurrency markets, understanding the thought process behind these studies.

Greene and Fielitz (1977):

This study primarily applied the rescaled range analysis method to examine the existence of long-term dependence in daily common stock price returns. The study used data from 200 common stocks listed on the New York Stock Exchange, specifically analyzing the daily composite returns from December 23, 1963, to November 29, 1968. Each stock had approximately 1220 data samples. The study estimated the Hurst value of the sample stocks using time lags of 10 and 50 units. The number of daily values of stocks with a lag of 10 and 50 units that exceeded 0.5 was 164 and 132, respectively. Most of the Hurst values were above 0.5, tentatively indicating positive dependence in the returns of these 200 stocks. However, when examining the Hurst value calculated with a lag of 50 units, approximately two thirds of the stocks had a lower Hurst value compared to the Hurst values calculated with a 10 unit lag. This outcome is as expected, where the increase in lags will decrease the Hurst value (Mandelbrot, 1971). Furthermore, the study calculated the first-order autocorrelation coefficients for these 200 stocks, and it was found that the absolute values of the coefficients for 160 stocks were below 0.1. These results collectively demonstrate the presence of positive and long-term dependence in the returns of these 200 stocks, thereby refuting the efficient market hypothesis (Hurst, 1951) ; (Mandelbrot, 1971).

Peters (1989):

This study also employed the rescaled range analysis method to explore the potential underlying structures in stock returns, bond returns, and the relative returns between stocks and bonds. The study examined the monthly returns of the S&P 500 from January, 1950 to June, 1988, along with 30-year treasury bonds, resulting in a total of 463 monthly data points. When compared to a random walk whose Hurst value is 0 and has no dependency on past values, the S&P 500 series had a Hurst value of 0.611 and a past value dependency of 0.168; the 30-year treasury bond series had a Hurst value of 0.614 and a past value dependency of 0.215; the relative returns between stocks and bonds series had a Hurst value of 0.658 and a past value dependency of 0.245. The number of years past for past effects to dissipate are 38 years, 38 tears, and 6 years for the S&P 500 series, the 30-year treasury bond series, and the relative returns between stocks and bonds series, respectively. The estimated Hurst exponent values for the capital market were all above 0.5, indicating that the behavior of these markets deviated from a random walk process.

Blank (1991):

This study primarily focused on investigating the presence of chaos in the futures market by utilizing the correlation dimension and Lyapunov exponent analysis methods. The S&P 500 index and soybean prices were chosen as the subjects of analysis. Due to the absence of reference statistics for both the correlation dimension analysis and the Lyapunov exponent analysis, the study used 100 sequences of randomly generated numbers from different distributions as comparative reference values for the mean correlation dimension analysis and Lyapunov exponent analysis of the sample sequences. The empirical results revealed that the correlation dimension values for the long-term S&P 500 index and soybean prices ranged from 2.3 to 2.7 and from 1.7 to 2.2, respectively; for the short-term S&P 500 index and soybean prices ranged from 1.5 to 1.7 and from 1.3. We can infer from this, as since the correlation dimension does not increase with the embedding dimension, all four of these futures support a non-linear process. Additionally, the Lyapunov exponent analysis values for both the S&P 500 index and soybean prices were both greater than zero, indicating the presence of chaos in these processes.

Hsieh (1991):

This study aims to investigate whether there is chaos in the volatility of stock returns from the beginning of 1963 to the end of 1987. The study initially conducted hypothesis testing assuming independent and identical distributions using the BDS statistic. The results provided sufficient evidence to reject the null hypothesis, indicating the presence of dependence among the data. However, rejecting the assumption of independent and identical distributions does not necessarily negate the efficient market hypothesis, as factors other than chaos can contribute to the observed dependence. Therefore, the study further analyzed factors such as non-stationarity, chaos, and nonlinear stochastic processes that could potentially contribute to the observed dependence. The findings indicated that the rejection of the assumption of independent and identical distributions was not due to policy changes or chaos but rather the presence of conditional heteroskedasticity in the data.

Tong and Chen (2022)

This study aims to analyze the price volatility of cryptocurrencies, particularly Bitcoin, exploring the impact of bullish and bearish information on cryptocurrency price fluctuations while also studying the long-term unpredictability and cyclical trends in cryptocurrency prices. It examines Bitcoin's daily closing prices from 2013 to 2021 and established that the price volatility of Bitcoin and other cryptocurrencies demonstrates non-linear dynamism, and does not conform to the random walk hypothesis. The study found that the fluctuations share a positive time correlation, with both optimistic (bullish) and pessimistic (bearish) information exerting approximately equal influence on price swings. Prices of cryptocurrencies reflect cyclical patterns, innate long-term unpredictability, and exhibit traits of fractals and chaos. The characteristics of clustering and persistence were noted, pointing to their inherent unpredictable and fractal traits in the long term. The investigation also highlighted a negligible leverage effect in Bitcoin, attributing price variations mainly to fundamental transactions and both bullish and bearish news. The dynamics of the price are complex and non-linear in nature. The study's findings serve as a warning to prospective investors to avoid speculative trades without a comprehensive grasp of the chaotic attributes of the cryptocurrency market. It also advises regulators to prudently handle cryptocurrency transactions and strive to comprehend the new attributes and laws of the market from a non-linear standpoint.

There are many other applications of chaos theory, and we have provided just a few relevant references that are closely related to this study. It is important to note that the results obtained may vary depending on the subject of analysis and the period studied.

3. Research Methods

In this chapter, we will systematically explain the various testing tools employed in this research, divided into five sections. The first section provides an introduction to the fundamental characteristics and functions of the statistical measures jointly proposed by Brock, Dechert, and Scheinkman. The second and third sections elucidate methods for testing stability and autocorrelation. The fourth section describes a tool for examining market efficiency, specifically the rescaled range analysis method. Lastly, the fifth section introduces a dimensional analysis approach for assessing data interrelationships, and extensively discusses the relationship between the results obtained from this method and the phenomenon of chaos.

Section 1.1: Brock – Dechert – Scheinkman test

The BDS (Brock, Dechert, Scheinkman) evaluation, titled after its authors, is a methodical technique to identify nonlinearity and dependence within a sequence of data points. In essence, it investigates whether a data series exhibits any characteristics of chaotic behavior. Typically, the test is applied to assess the proposition that the residuals from an ARMA (Autoregressive Moving Average) model adhere to the independent and identically distributed (i.i.d.) principle.

The null hypothesis of the BDS test is that the time series is independent and identically distributed. If we reject the null hypothesis, this suggests that there is some form of structure in the data - this could be linear or nonlinear dependence.

The BDS test is a form of a correlation integral test. It compares the likelihood of pairs of points being close in the m-dimensional space at different embedding dimensions. If the data is independent, the proportion of close points should not change with the embedding dimension.

To compute for the BDS test, we first define vectors X_t^m and X_s^m with observations of T and a given embedding dimension of m:

$$X_t^m = (X_t , X_{t+1} , . . . , X_{t+m-1})$$

$$X_s^m = (X_s , X_{s+1} , . . . , X_{s+m-1})$$

We then compute the correlation sum evaluated as the percentage of all vector pairs (X_t^m , X_s^m) which exist within an ε distance of each other with the following equation:

$$C_{m,T} (\varepsilon) = \sum_{t<s} I\varepsilon \, (X_t^m , X_s^m) * 2 / (T_m (T_m - 1))$$

In this formula, $I\varepsilon$ (X_t^m , X_s^m) equals 1 when $\|X_t^m - X_s^m\| < \varepsilon$, and 0 otherwise. $\|X_t^m - X_s^m\|$ denotes the Euclidean norm.

According to equation above, the overall dimension is a measure of the extent to which the distances between data sets are less than a certain value ε. Therefore, the choice of m and ε values will affect the magnitude of the overall dimension. For a given m value, a smaller ε value corresponds to a smaller number of data points falling within this interval distance, and vice versa (Hsieh, 1989). defines the ε value as a multiple of the standard deviation of the sequential data. The standard deviation multiples he used are 1.50, 1.25, 1.00, 0.75, and 0.50. This study will refer to Hsieh's standard deviation multiples as the criterion for selecting the ε value.

The BDS statistic is calculated as the difference in the natural log of the correlation sums for embedding dimensions m and m-1, multiplied by the square root of the number of observations. The equation of which looks like this:

$$BDS = \sqrt{N} * (\log(C(m , \varepsilon)) - \log(C(m-1 , \varepsilon)))$$

If the time series is i.i.d., then the BDS statistic has an asymptotic standard normal distribution under certain conditions. This is used to form a test of the null hypothesis of i.i.d.

Brock et al. (1987) applied Monte Carlo simulation and found that sample distribution can provide a better approximation to population distribution under the following conditions: (1) when the data size is equal to or greater than 500 observations, (2) when the dimension does not exceed 5, and (3) when the value is within the range of 0.5 to 2.0 times the standard deviation. Additionally, BDS exhibits increasing test power towards 100% as the sample size increases for seven linear and nonlinear models, namely AR(1), MA(1), tent map, threshold autoregression, nonlinear moving average, autoregressive conditional heteroscedasticity (ARCH), and generalized autoregressive conditional heteroscedasticity (GARCH). Brock et al. further suggest that except for ARCH and GARCH models, BDS maintains high testing power regardless of whether the data used are the original series or the residual series. Moreover, Hsieh (1991) demonstrates that the BDS test possesses good testing power for rejecting the i.i.d assumption in the following four types of possible characteristics: linear dependence, non-stationarity, chaos, and nonlinear stochastic process.

Therefore, to test for the presence of nonlinear relationships in the data, Baumol and Benhabib (1989) suggest pre-whitening the data by estimating the residuals of a linear model before conducting the BDS test. If the BDS statistic rejects the i.i.d. assumption, it indicates the presence of nonlinear phenomena in the data.

Regarding nonlinear testing, there are other methods such as Tsay (1986) test and Engle (1982) test. Tsay's test exhibits worse results than the BDS test and Engle's test for ARCH and GARCH models. Engle's test on the other hand is more suited for testing ARCH models. The BDS test yields better results for general types of nonlinear models. Hence, in this study, the BDS test will be used to examine the presence of nonlinear relationships in the data.

Section 3.2: Rescaled Range (R/S) method for the Hurst Exponent

Through extensive research on many time series data, it has been observed that higher values in the early period tend to be followed by higher values in the later period (Mandelbrot & Hallis, 1969), as seen in phenomena such as weather data (Hurts, 1951) and futures contract prices (Helns et al., 1984). This phenomenon is commonly referred to as persistence. However, persistence is not caused by general sequence correlation but rather by a nonlinear phenomenon characterized as "long-term statistical dependence with infinite memory." This concept is indicative of the presence of nonlinear phenomena (Mandelbrot, 1972).

The efficiency of a market and the existence of persistence in market volatility are directly related. Previous studies generally assume that stock returns exhibit finite memory effects, where dependence decreases rapidly as the time lag increases. Under this assumption, coupled with the fact that the autocorrelation coefficients of the first several lags are very close to zero, the volatility process of the stock market is approximated as a random walk (Fama, 1970).

To obtain the Hurst exponent through the rescaled range method, we first set a stationary timeseries of T observations as such:

$$X_t = (X_t , X_{t+1} , . . . , X_T)$$

Then we calculate the partial sum of the first d deviations from the mean can be given by the following equation

$$R(t , d) = \max \{X^*(t + u) – X^* (t) – (u / d) [X^*(t + d) – X^*(t)]\} - \min \{X^*(t + u) – X^* (t) – (u / d) [X^*(t + d) – X^*(t)]\}$$

Then we compute the standard deviation, which is the S component of the rescaled range, with the following equation:

$$S^2 (t , d) = [\sum_{u=1}^{d} X^2 (t + u) / d] – [\sum_{u=1}^{d} X (t + u) / d]^2$$

And thus, we get:

$$R/S = R(t , d) / S(t , d) , \quad 3 \leq d \leq t$$

In the research conducted by Wallis and Matalas (1970), two feasible methods were proposed for dealing with time lags and starting periods: F Hurst and G Hurst. F Hurst calculates the R/S values for all combinations of different starting points and time lags, while G Hurst calculates the R/S values for a specific time lag and starting period of no more than 15 observations. The study indicated that although F Hurst provides smaller biases than G Hurst, using G Hurst helps reduce computational complexity. When the sample size is 1000 and the minimum time lag is 50 observations, the calculated H value using G Hurst (GH50) is biased by only approximately 0.02 on the upper side. Furthermore, the bias decreases as the sample size increases. In this study, considering the total of 8760 observations in the dataset, we follow the G Hurst method in choosing the time lag and starting period to reduce computational complexity. It is anticipated that using G Hurst to calculate the Hurst value will result in a bias smaller than 0.02 (Mandelbrot, 1972).

The implications of the Hurst value can be summarized into the following:

1. When the Hurst value = 0.5, the series at hand shows it to be a Geometric Brownian Motion or has finite memory.
2. When the Hurst value > 0.5, the series at hand shows trending behavior with positive long-term dependence, or in other words, persistence.
3. When the Hurst value < 0.5, the series at hand shows signs of either being mean reverting or stationary, with negative long-term dependence.

Section 3.3: Correlation Dimension

In section 1, we mentioned that the overall dimension is used to measure the degree to which distances between data sets are smaller than a certain value. In this section, we will extend the concept of overall dimension and briefly explain the calculation process of correlation dimension, as well as the relationship between correlation dimension and chaotic phenomena.

Grassberger and Procaccia (1983) define the correlation dimension of a sequence $\{X_t\}$ as "the general logarithm of the overall dimension divided by the general logarithm of ε," which in equation is as the following:

$$V_m = \lim_{\varepsilon \to 0} [\ \log C_{m,T}(\varepsilon) / \log(\varepsilon)\ ]$$

where V_m represents the correlation dimension of an m dimensional vector. Brock and Sayers (1988) state that the term "dimension" is used to measure the complexity of a system. Therefore, the correlation dimension can be seen as a measure of the system's degrees of freedom. For example, the dimension of a point is zero, the dimension of a line is one, and the dimension of a cube is three. In the case of completely unordered white noise processes, the dimension of which would be infinite. Thus, if the V_m values of a time series do not increase with the increase in m, indicating the existence of a limiting value V for V_m, then it means that the series can be measured with fewer variables (approximately equal to V), and it may exhibit chaotic phenomena. A chaotic system must have a positive and finite V (Frank et al., 1988). However, Casdagli (1991) points out that a high-dimensional deterministic system is indistinguishable from a random system.

Other common detection tools for assessing whether a sequence exhibits chaotic phenomena, besides the aforementioned correlation dimension analysis, include the Lyapunov exponent (Holf et al., 1985) and the Kolmogorov entropy (Grassberger & Procaccia, 1983). However, these methods require prior knowledge or assumptions about the functional form of the sequence's trajectory, making practical calculations difficult. Therefore, this study employs the relatively simpler method of the correlation dimension analysis.

4. Empirical Analysis:

In this chapter, several testing methods mentioned in the previous chapter will be applied to sequentially examine whether the behavior of Bitcoin volatility can be explained by a certain number of variables. The entire analytical framework is based on the BDS statistic, with separate sections dedicated to each aspect. In Section 1, the BDS statistic is first employed to test whether the raw data of the hourly Bitcoin closing price series adheres to the assumption of independent identically distributed. If the null hypothesis is rejected, as suggested by Hsieh (1991), this phenomenon could be caused by four processes: nonstationary, linear dependence, chaos, and nonlinear stochastic processes. Therefore, in the following sections, this study will conduct corresponding tests for the aforementioned four potential factors in the hourly Bitcoin closing price.

Section 4.1: Testing for Independent Identically Distributed

This study tests the hourly closing prices of Bitcoin starting from January 1, 00:00, 2021 to January 1, 00:00, 2022. First, a test is conducted for the independent identically distributed property of the raw data. Let the hourly closing prices of Bitcoin be denoted as series $\{X_t\}$, where (X_1, X_2, ..., X_{8760}), X_1 represents the first recorded closing price on January 1, 00:00, 2021, and X_{8760} represents the closing price on January 1, 00:00, 2022, which is 46650.66 USD. As for the selection of the ε value, we follow the recommendations of Brock et al. (1987), which are 0.5, 1.0, 1.5, and 2.0 times the standard deviation of the series. The dimension m ranges from 2 to 5. Under the null hypothesis H_0, the sequence is independent and identically distributed, while under the alternative hypothesis H_1, the sequence is non iid. The BDS statistic is calculated based on equations mentioned earlier to test the hypothesis, and the results are presented in Table 1.

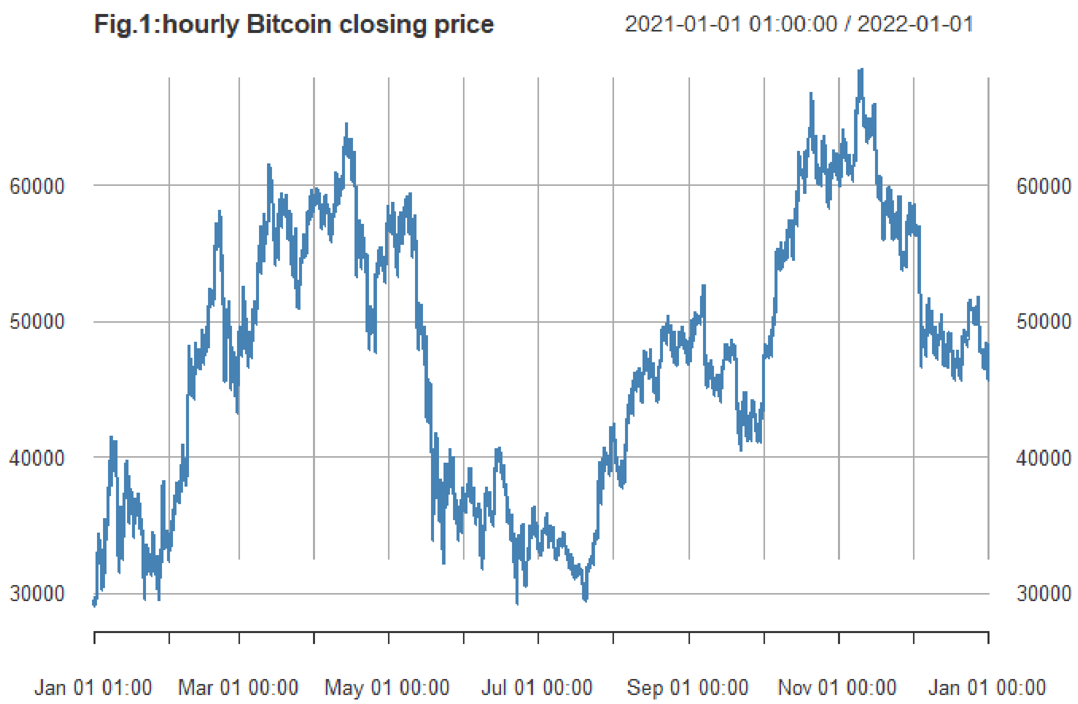
In Table 1, for each selected m value and ε value, a corresponding BDS value is determined. A larger BDS value indicates a higher likelihood of rejecting the iid hypothesis. Since the BDS statistic approximates a standard normal distribution, if the absolute value of the BDS value exceeds 1.96 at a significance level of 5%, it indicates that the null hypothesis of independent identically distributed is rejected, implying that the sequence is non iid. Except for the last column, the BDS values generally increase with the increase of embedded dimension. This demonstrates a stronger dependence between consecutive trading data of longer hours compared to consecutive closing prices of shorter hours. Moreover, when the dimension is fixed, the BDS values decrease with the increase of ε, which is consistent with theoretical expectations. Overall, since all the BDS values in the table are much larger than 1.96, and the p values are all less than 0.05, rejecting the null hypothesis, there is sufficient evidence to confirm that the sequence is non iid. In other words, the hourly Bitcoin closing price series is not white noise, indicating some underlying characteristics among the data. Based on the properties of BDS, it is known that these characteristics could be one of the following: linear dependence, non-stationarity, chaos, or nonlinear stochastic processes. In the next section, this research will first investigate the stationarity of the series $\{X_t\}$ to understand whether rejecting the iid hypothesis is due to the presence of non-stationarity.

```
         BDS Test

data:  BTCdf

Embedding dimension =  2 3 4 5

Epsilon for close points =   4906.419  9812.837 14719.256 19625.675

Standard Normal =
      [ 4906.4187 ] [ 9812.8373 ] [ 14719.256 ] [ 19625.6747 ]
[ 2 ]      3321.283      968.8253      459.3773        355.0555
[ 3 ]      6764.509     1309.1459      502.4351        351.4525
[ 4 ]     15279.335     1840.3196      556.1381        349.2016
[ 5 ]     37824.793     2719.2826      631.9325        353.3801

p-value =
      [ 4906.4187 ] [ 9812.8373 ] [ 14719.256 ] [ 19625.6747 ]
[ 2 ]             0             0             0              0
[ 3 ]             0             0             0              0
[ 4 ]             0             0             0              0
[ 5 ]             0             0             0              0
```

Table 1 BDS test statistics of the hourly Bitcoin closing price series

Section 4.2: Testing for Non-stationarity

In the field of finance and economics, non-stationarity refers to changes in the structure of a sequence caused by various factors such as technological or financial innovations, or changes in government policies. In the year 2021, the cryptocurrency market experienced massive external shocks, raising doubts on whether the sequence structure could remain unchanged. Observing the trend chart of the hourly Bitcoin closing price series for this period (Fig.1), it can be seen that the whole series exhibits significant fluctuations. Additionally, the autocorrelation function of this sequence (see Fig.2) shows a slow decay, indicating that this time series is non-stationary. This study utilizes the Augmented Dickey-Fuller unit root test to examine the non-stationarity of the time series. The test result shows that the p value is 0.5055 and the F-value is -2.172, which is smaller than the critical value of 6.25. This indicates that the null hypothesis cannot be rejected, implying that time trend contributes to the non-stationarity of the Bitcoin closing price index, aligning with general expectations.



To investigate whether non-stationarity is the main factor causing the rejection of the independent identically distributed hypothesis, the data is transformed appropriately to satisfy stationarity before conducting the BDS test again, and then the results of which are compared to the BDS values of the original time series. We transform the data by taking the first-order difference of the data points in the hourly Bitcoin closing price series. The subsequent output series is the hourly logarithmic return rate of Bitcoin. Fig.3 is the series diagram of the Bitcoin return series. We can see from the graph that there are large instances of fluctuations within the series, while also these fluctuations are rather clustered. Again, we test this new series to check it the non-stationarity has been accounted for after the transformation. The Augmented Dickey Fuller test and the Phillip Perron test are used. Both tests return p values of less than 0.05, indicating that the Bitcoin return series is indeed stationary, and the non-stationary characteristics have been accounted for through the transformation. We take the Bitcoin return series and perform the BDS test on it. The ε value is set to 0.5, 1.0, 1.5, and 2.0, and the m value is still selected as 2 to 5. The results are shown in Table 2. Comparing Table 1 and Table 2, all the BDS values in Table 2 are lower than those in Table 1, indicating that the data, after being transformed by taking the natural logarithm and then first-order differencing, can explain some of the non-iid characteristics. However, the absolute values of the BDS values in Table 2 are still larger than the critical value of 1.96, with p values all less than 0.05, rejecting the null hypothesis, which suggests that non-stationarity is not the main factor causing non-iid. In the next section, this study will explore another possible factor contributing to non-iid: linear dependency.

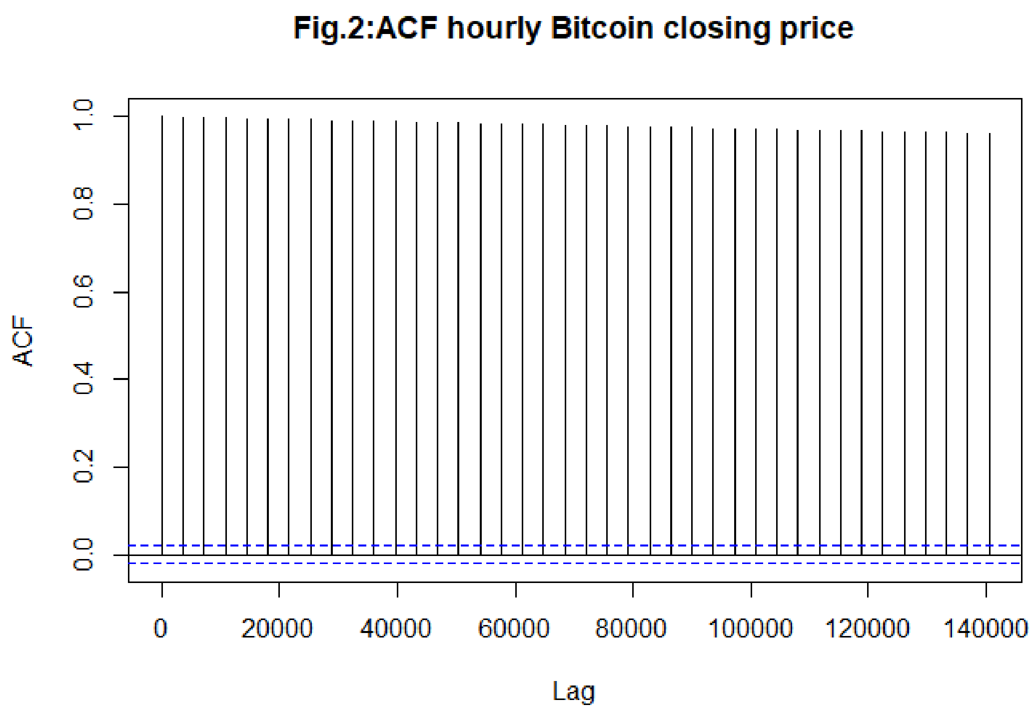


Fig.2:ACF hourly Bitcoin closing price
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
20000
40000
60000
80000
100000
120000
140000
Lag

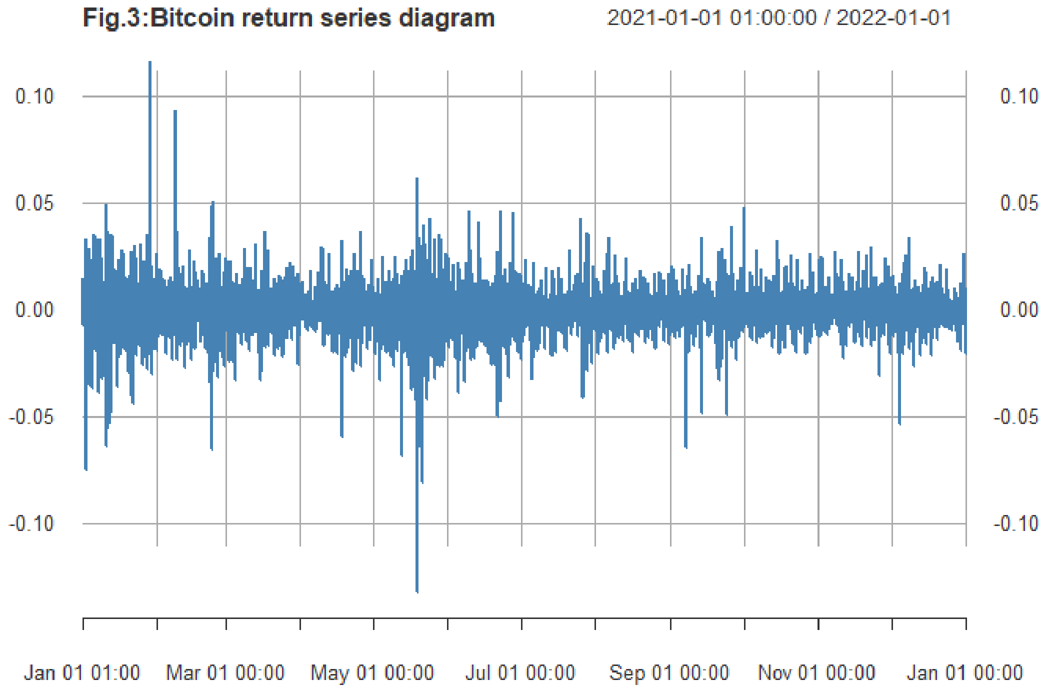


Fig.3:Bitcoin return series diagram
2021-01-01 01:00:00 / 2022-01-01
0.10
0.05
0.00
-0.05
-0.10
Jan 01 01:00
Mar 01 00:00
May 01 00:00
Jul 01 00:00
Sep 01 00:00
Nov 01 00:00
Jan 01 00:00

```
        BDS Test

data:  BTCts

Embedding dimension =  2 3 4 5

Epsilon for close points =  0.0047 0.0094 0.0140 0.0187

Standard Normal =
      [ 0.0047 ] [ 0.0094 ] [ 0.014 ] [ 0.0187 ]
[ 2 ]    19.2800    20.0350   18.5583    16.6523
[ 3 ]    24.9505    24.3766   21.9099    19.4250
[ 4 ]    29.3651    27.4017   24.3332    21.7747
[ 5 ]    33.6647    30.0168   26.1618    23.4777

p-value =
      [ 0.0047 ] [ 0.0094 ] [ 0.014 ] [ 0.0187 ]
[ 2 ]          0          0         0          0
[ 3 ]          0          0         0          0
[ 4 ]          0          0         0          0
[ 5 ]          0          0         0          0
```

Table 2 BDS test statistics of the Bitcoin return series

Section 4.3: Autocorrelation, Linear Dependency, and Heteroskedasticity

The focus of this section is on the newly generated series that takes the hourly Bitcoin closing price data and applies natural logarithmic transformation and first-order differencing to form a new sequence. The autocorrelation coefficient of the sequence is calculated, and its null hypothesis is tested. A suitable time series model is then fitted to eliminate linear dependencies in the sequence, and the resulting residuals form a new time series. Subsequently, whether the residual series are iid or not is tested.

The underlying characteristics can be roughly categorized into either linear or nonlinear characteristics. Therefore, this study aims to identify a linear model to account for the linear dependency of this stationary time series. When fitting a time series ARMA (p, q) model, the values of p and q are determined using the Box-Jenkins methodology, based on the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the sequence. In Fig.4, the autocorrelation function converges gradually after lag 4, while Fig.5 shows an extremely jagged partial autocorrelation function. Taking both figures into account, the preliminary estimation suggests $p \approx 4$ and $q \approx 4$.

In fitting the ARMA model, this study attempted several models with a loop in Rstudio for p and q values both not exceeding 5. While it is true that the acf seems to have slightly above reasonable number of significant lags at higher levels, and the pacf clearly shows many significant lags at higher orders and seems to be converging at a very slow lag, we do not see any noticeable seasonality within the two graphs. We also constrict the maximum number of both the p and q to relax computational stress. The results from the Rstudio loop indicate that the AR(4) process has both the lowest AIC and the lowest BIC. Hence, we fit the AR(4) model to the hourly Bitcoin return series, allowing its residuals sequence to be examined for linear dependence, determining if it contributes to the rejection of the independent and identically distributed assumption. Table 4-7 shows the test results of the BDS values on the residual terms after fitting the AR(4) model to the Bitcoin return series. Comparing Table 4-3 and Table 4-7, the BDS values of the fitted AR(4) model are smaller than those of the Bitcoin return series before fitting the model, indicating that linear dependence partially explains the non-i.i.d. outcome of the BDS test. Even though the BDS test

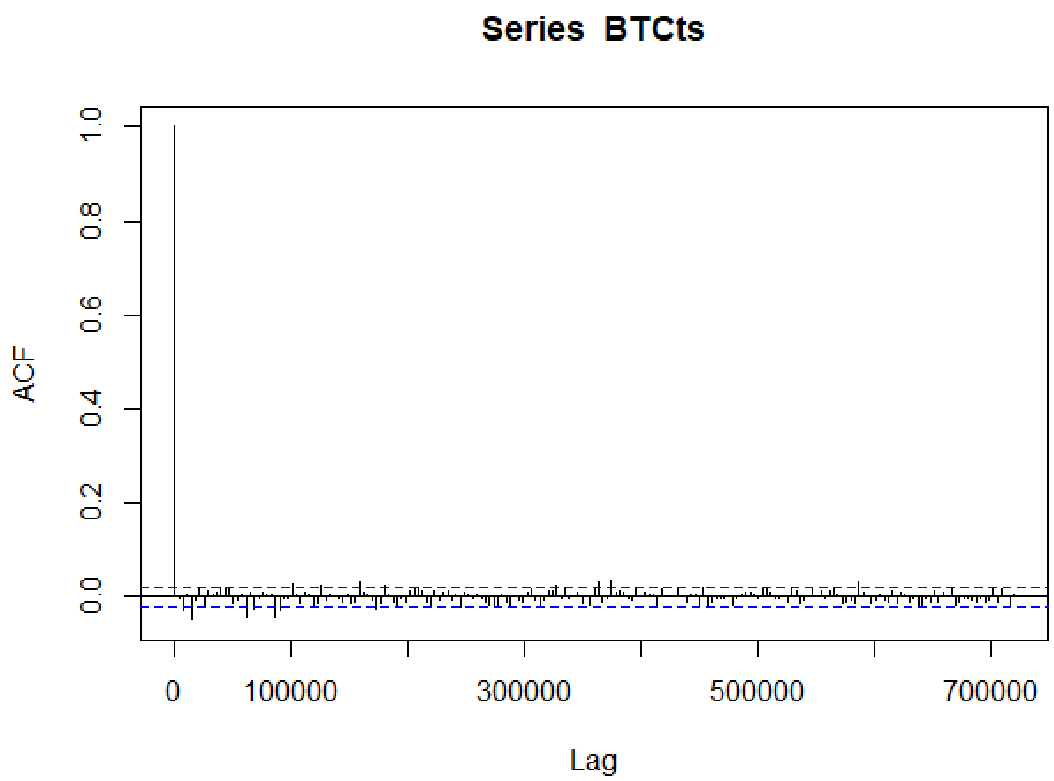

Series BTCts
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
100000
300000
500000
700000
Lag

Fig.4

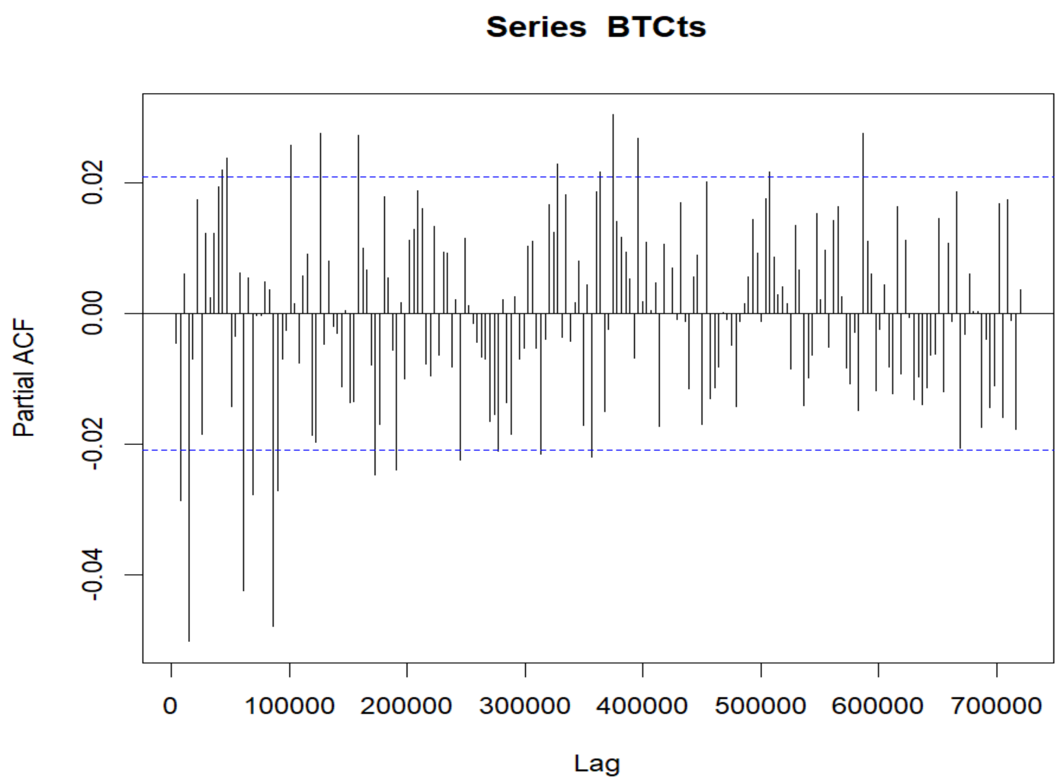

Series BTCts
Partial ACF
0.02
0.00
-0.02
-0.04
0
100000
200000
300000
400000
500000
600000
700000
Lag

Fig.5

results on the residuals of the fitted AR(4) are lower than before fitting, the difference between fitted and non-fitted results are relatively small. Also, the observed absolute values of the BDS values in Table 3 are larger than the critical value of 1.96 for a standard normal distribution at a 5% significance level, the p values are also less than 5%, rejecting the null hypothesis. Therefore, we still reject the independent and identically distributed hypothesis. We also conduct tests to see if the AR(4) is a good fit. The acf plot of the residuals of the shows no sign of autocorrelation, yet the acf plot of the residuals squared show clear terms of volatility, being correlated at most lags.

```
                BDS Test

 data:  BTCarima$residuals

 Embedding dimension =  2 3 4 5

 Epsilon for close points =  0.0047 0.0093 0.0140 0.0187

 Standard Normal =
       [ 0.0047 ] [ 0.0093 ] [ 0.014 ] [ 0.0187 ]
[ 2 ]     18.9796    19.7893   18.2848    16.3789
[ 3 ]     24.5300    24.1073   21.6109    19.1008
[ 4 ]     28.9189    27.0532   23.9533    21.3778
[ 5 ]     33.0524    29.5106   25.6627    23.0157

 p-value =
       [ 0.0047 ] [ 0.0093 ] [ 0.014 ] [ 0.0187 ]
[ 2 ]           0          0         0          0
[ 3 ]           0          0         0          0
[ 4 ]           0          0         0          0
[ 5 ]           0          0         0          0
```

Table 3

To address this issue of heteroskedasticity, a GARH model is then fitted to the residuals of the AR(4). Four different GARCH(p, q) models were constructed, being (1, 1), (1, 2), (2, 1), and (2, 2). The best model is then selected through comparing the AIC and the BIC. The optimal GARCH model selected is the GARCH(1, 2), with an AIC of -6.7250 and a BIC of -6.718, the lowest out of the four. From the test results of the Ljung-box test on the standardized residuals, the p value for all three lags tested are greater than 5%, failing to reject the null hypothesis, indicating that there is no significant evidence of residual autocorrelation in the data and that the residuals are independent and not correlated. Test results of the Ljung-box test on the standardized squared residuals also show that for all three lags, the p values are greater than 5%, failing to reject the null hypotheses, indicating that there is no significant clustering or pattern in the squared residuals, the assumption of no autocorrelation in the conditional variance is reasonable, and that the model captures the dynamics of volatility. For the last test, the ARCH LM test, all three lags have p values of greater than 5%, failing to reject the null hypothesis, indicating that no significant evidence of ARCH effects or conditional heteroscedasticity remain in the data and that the model captures these characteristics well. These properties indicate a good fitting model. The BDS test is then ran on the residuals and the volatility of the fitted GARCH(1, 2) model. We still select the same parameters with ε value set to 0.5, 1.0, 1.5, and 2.0, and the m value still selected as 2 to 5. The BDS test results for both the residuals (Table 4) and the volatility (Table 5) show that they are non-iid. The results of this test implies that the residuals series and the volatility series should be assessed with nonlinear models.

```
        BDS Test

data:  BTCgarch@fit$residuals

Embedding dimension =  2 3 4 5

Epsilon for close points =  0.0047 0.0094 0.0140 0.0187

Standard Normal =
      [ 0.0047 ] [ 0.0094 ] [ 0.014 ] [ 0.0187 ]
[ 2 ]    19.3671    20.0608   18.4332    16.4789
[ 3 ]    24.8743    24.2819   21.6671    19.1778
[ 4 ]    29.1815    27.2039   24.0303    21.4857
[ 5 ]    33.3042    29.6915   25.7845    23.1635

p-value =
      [ 0.0047 ] [ 0.0094 ] [ 0.014 ] [ 0.0187 ]
[ 2 ]          0          0         0          0
[ 3 ]          0          0         0          0
[ 4 ]          0          0         0          0
[ 5 ]          0          0         0          0
```

Table 4

```
        BDS Test

data:  BTCgarch@fit$sigma

Embedding dimension =  2 3 4 5

Epsilon for close points =  0.0017 0.0033 0.0050 0.0066

Standard Normal =
      [ 0.0017 ] [ 0.0033 ] [ 0.005 ] [ 0.0066 ]
[ 2 ]   251.2741   168.8906  142.9181   132.8401
[ 3 ]   344.1667   181.7647  141.4085   126.5508
[ 4 ]   491.1679   197.5975  140.3590   120.6759
[ 5 ]   738.4623   220.2116  141.8706   116.8920

p-value =
      [ 0.0017 ] [ 0.0033 ] [ 0.005 ] [ 0.0066 ]
[ 2 ]          0          0         0          0
[ 3 ]          0          0         0          0
[ 4 ]          0          0         0          0
[ 5 ]          0          0         0          0
```

Table 5

Section 4.4: Confirmation of nonlinearity

After testing for non-stationarity and linear dependency, the results from the BDS test indicate that the both the residuals series of the fitted GARCH(1, 2) and the volatility series of the fitted GARCH(1, 2) are likely to have nonlinear characteristics. Therefore, this study employs the Rescaled Range (R/S) analysis method to confirm this finding. To get the Hurst exponent, estimating the value of R/S for each day requires calculating the corresponding R/S for different starting points and time lags. In this study, G Hurst with smallest lags at GH(20) is primarily used, divided into 30 blocks to calculate the R/S values. The average of the R/S values for each lag difference is then obtained, and the corresponding dimensions are determined.

The Hurst exponent for the residuals series of the fitted GARCH(1, 2) is 0.5318611, and the Hurst exponent of the volatility series of the fitted GARCH(1, 2) is 0.8272923. Additionally, based on the research of Mandelbrot (1972), the data correlation also denoted as Cr, is calculated to be approximately 0.04 and 0.58. Unlike sample autocorrelation, data correlation is different because the autocorrelation function assumes that the distribution of the sample follows a Gaussian distribution, resulting in better measurement of short-term dependence but often underestimating long-term dependence for non-Gaussian distributions. Considering the above results, it can be determined that the stationary series exhibits positive long-term dependence and does not follow a Gaussian distribution and are not random walks. Moreover, due to its persistence, the sequence is confirmed to be nonlinear, contradicting the efficient market hypothesis.

Section 4.5: Determining Chaos Characteristics

In this part of the study of time series data, it can be tentatively concluded that Bitcoin time series shows the characteristics of a complex and nonlinear system. I would like to further expand on the idea that perhaps the Bitcoin time series has chaotic properties. Hence, the incorporation of methodologies derived from chaos theory, including phase space reconstruction of both the residuals series of the fitted GARCH(1, 2) and the volatility series of the fitted GARCH(1, 2), calculating the related Takens' vectors, and the calculation of the correlation dimension to determine if the Bitcoin has a law that governs the underlying system.

Takens' vectors, or delay vectors, are used to populate the reconstructed phase space. Each vector represents a distinct state of the system at a given point in time. Takens' vectors serve as a bridge between the time series and its corresponding dynamical system, revealing a geometric picture of system dynamics that can assist in identifying signatures of chaos. They are constructed by delaying the original time series, effectively capturing the evolution of the system across different states.

To build a Takens vector, the time delay denoted as τ, and the embedded dimension denoted as d are the utmost important two parameters to estimate. At its most fundamental, the Average Mutual Information (AMI) measures the volume of information two random variables mutually contain. These variables correspond to the initial time series and its counterpart that has been shifted by varying τ levels. As such, the point where the function stabilizes or initially intersects zero provides an accurate gauge of the time delay retaining the minimal data on the current time point. This approach enjoys widespread use given its capacity to accommodate nonlinear correlations. The Average Mutual Information (AMI) for both the residuals series of the fitted GARCH(1, 2) and the volatility series of the fitted GARCH(1, 2) are shown in Fig. 6 and Fig.7 . The τ for the residuals series would thus be 1, and the τ for the volatility series would be 8. We next look for the optimal embedding dimension with Cao's method. Cao's method does not rely on parameters other than the embedded time delay as calculated from the average moving information method. From Fig.8 and Fig.9, The d for the residuals series would thus be 12, and the d for the volatility series would be 8. With these parameters, we can construct the Taken's vector and plot the reconstructed phase space diagram. Fig.10 and Fig.11 show the reconstructed phase space plotted on two dimensions, the shape of which are both complex, indicating the possibilities of strange attractors.

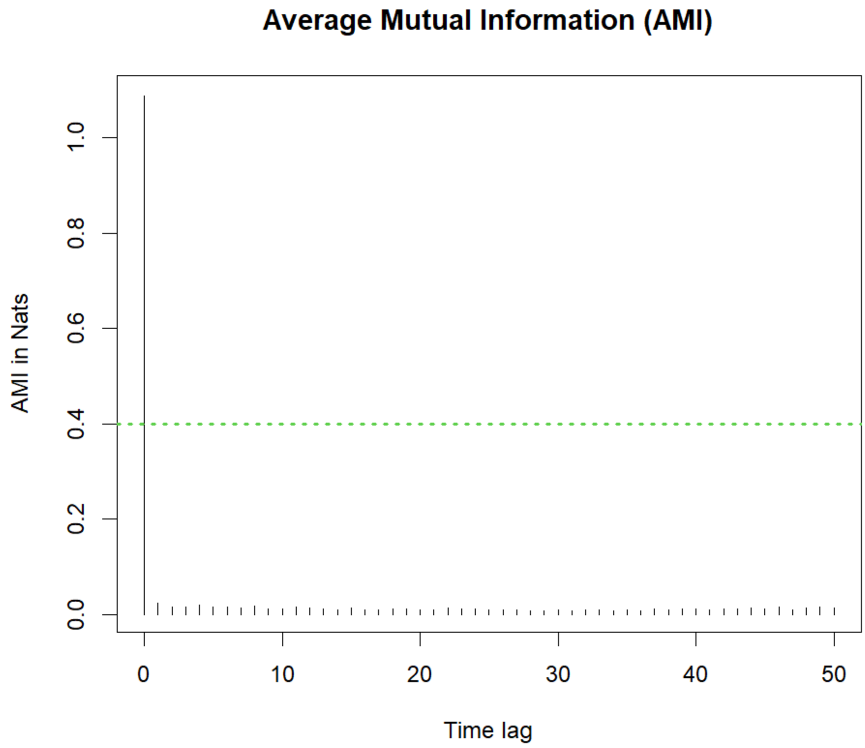


Fig. 6

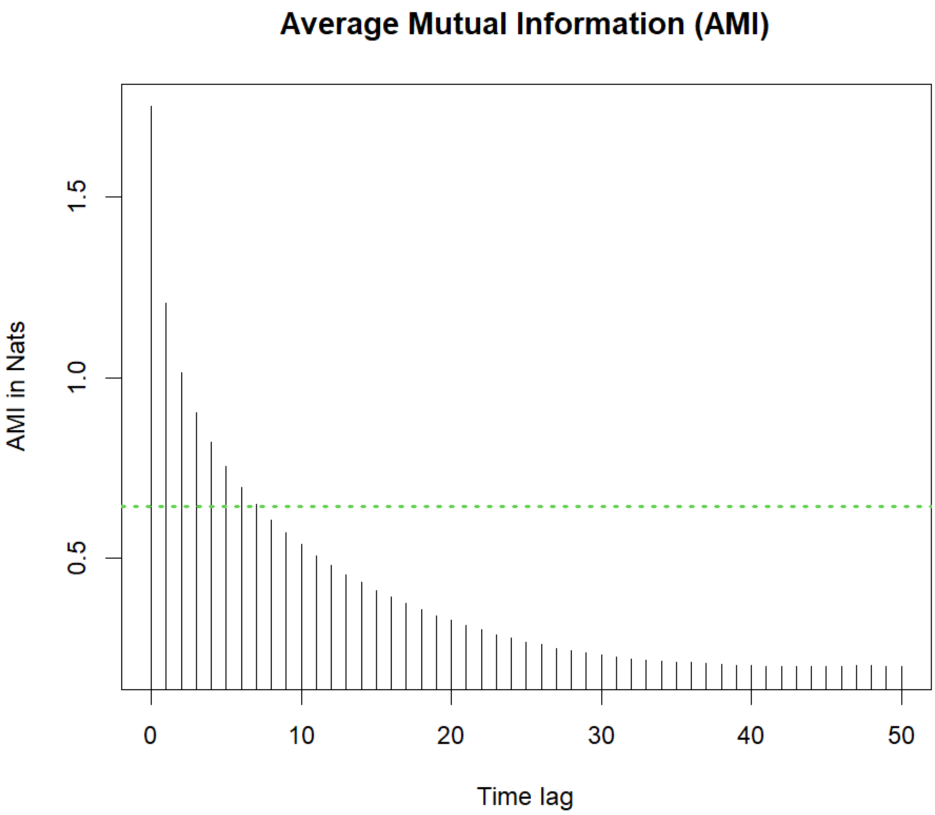


Fig.7

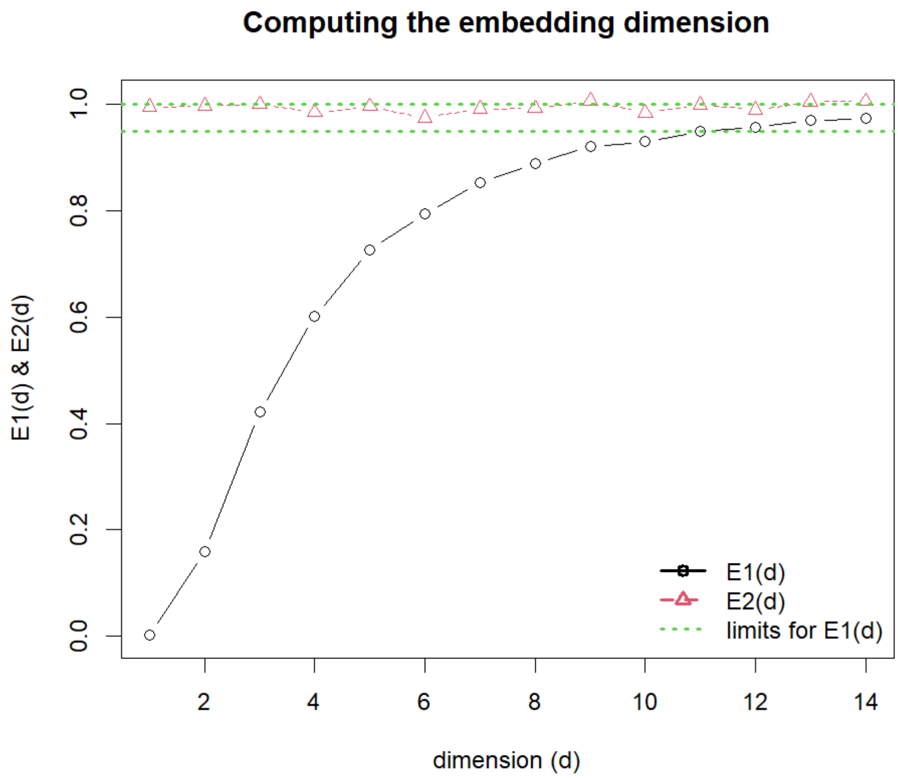


Fig.8

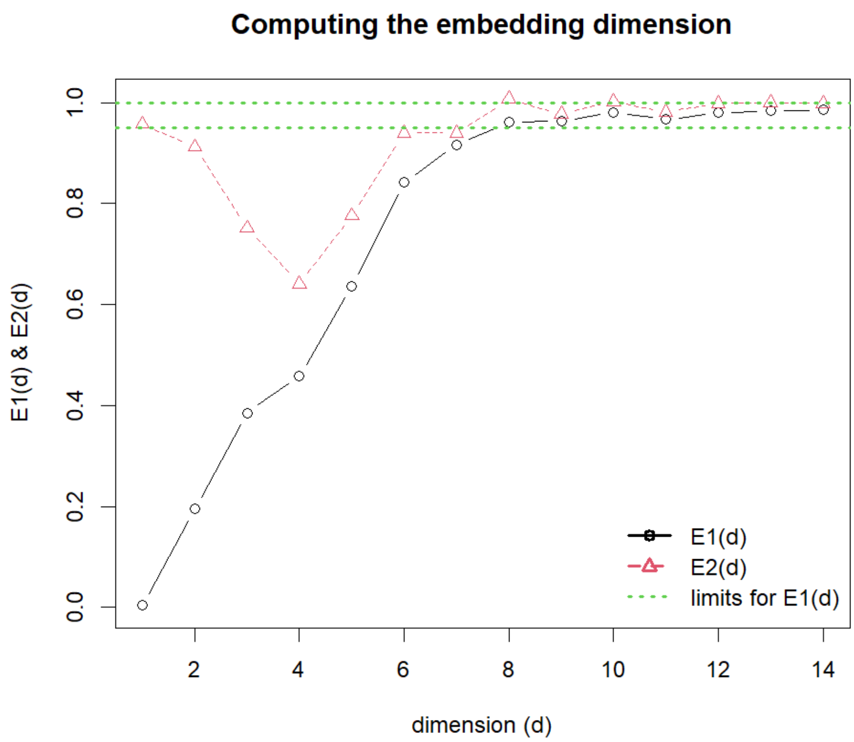


Fig.9

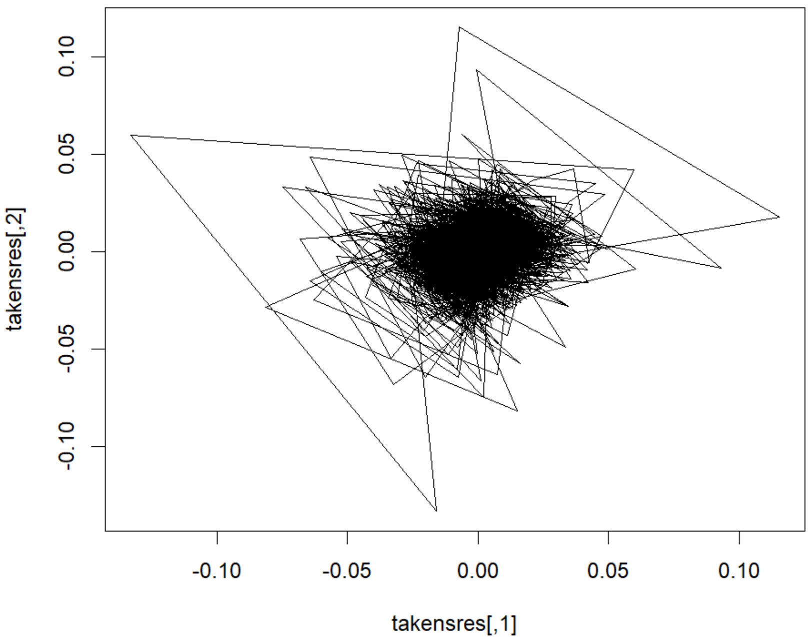

takensres[,2]
takensres[,1]
0.10
0.05
0.00
-0.05
-0.10
-0.10
-0.05
0.00
0.05
0.10

Fig.10

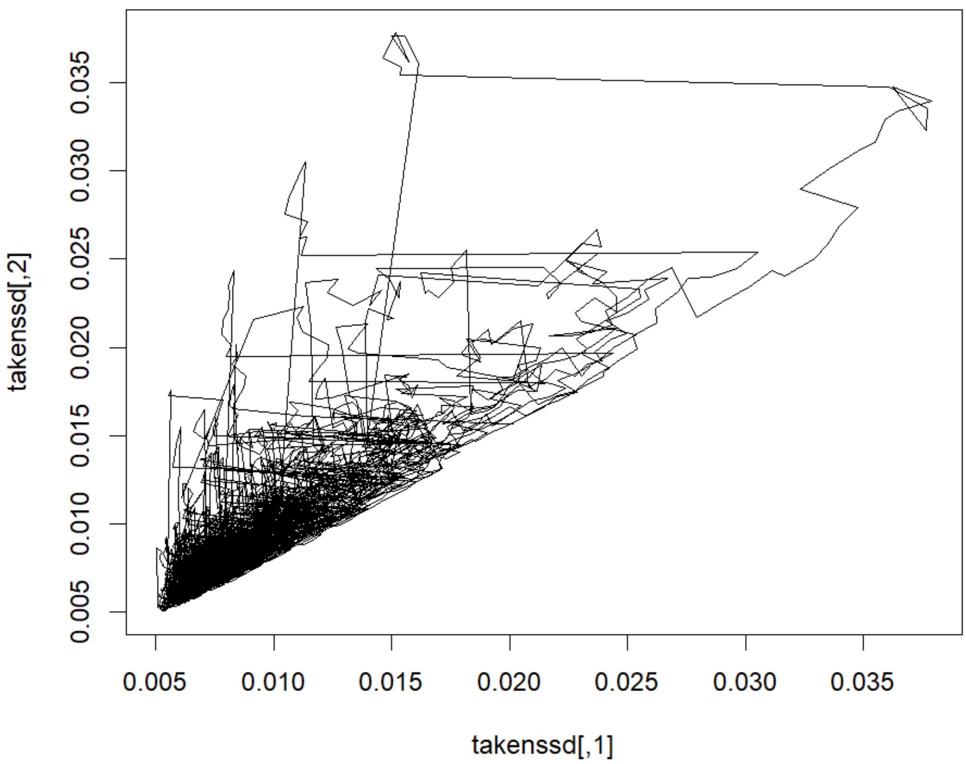

takenssd[,2]
takenssd[,1]
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.005
0.010
0.015
0.020
0.025
0.030
0.035

Fig.11

Now that both the time delay and the optimal embedded dimension are known, I can calculate the correlation dimension of the reconstructed phase space. The correlation dimension is a quantitative measure of the complexity of a dynamical system, representing the rate of increase of new states or points in phase space as the scale of observation changes. For the maximum embedded dimension, I tried all combinations until the computer could not estimate the number, which was 102 dimensions for the residual series (Fig.12), and 144 dimensions for the volatility series (Fig.). The results conflict with the rules a for a chaotic system, in which the correlation dimension had to be converging. While the residual series does seem to be slowly declining the point at where the correlation dimension plateau's is beyond the boundaries. The sharp drops in the figures are also something to note. This may be due to the complexity of the system, noisiness of the data, and perhaps and inadequate approach to the system.

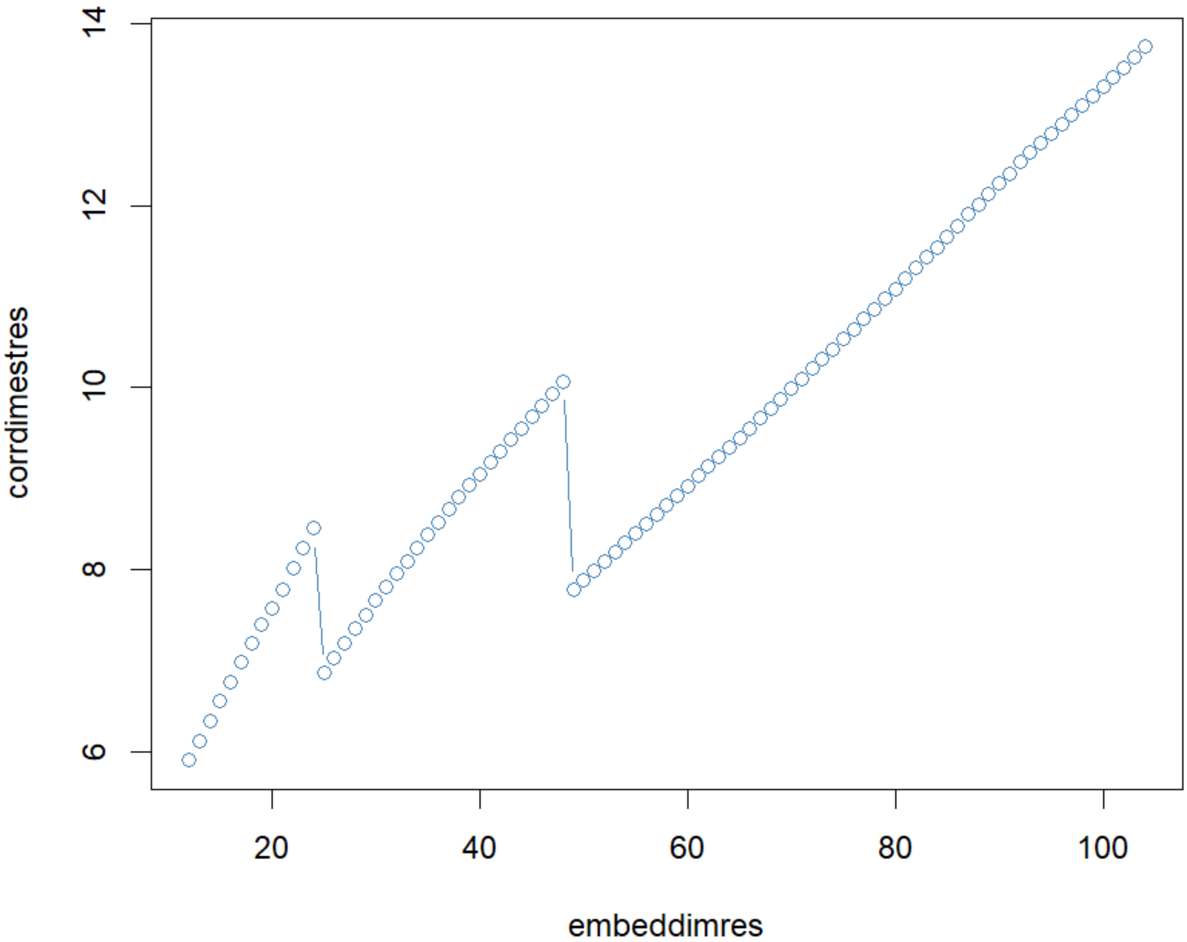


Fig.12

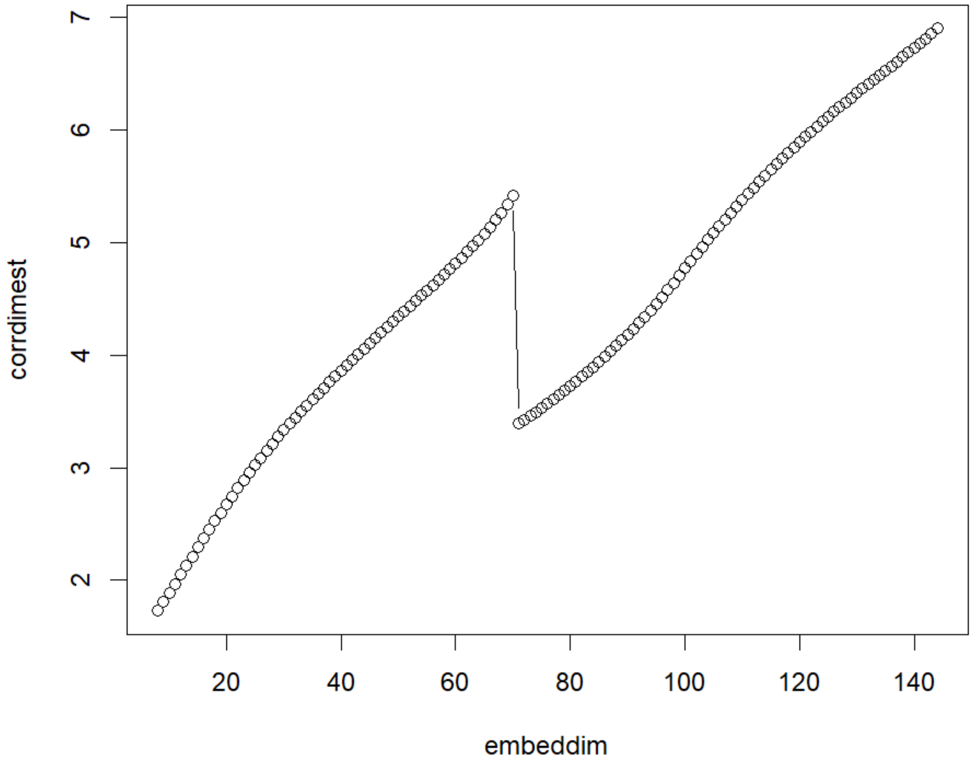


Fig.13

5. Conclusion:

This study examines whether the hourly closing data of Bitcoin from January 2021 to December 2021 follows an independent and identically distributed (iid) distribution. The results indicate that there is sufficient evidence to reject the hypothesis of independent and identically distributed behavior of the Bitcoin time series at a significance level of 5%. However, rejecting the hypothesis of independent and identically distributed behavior does not necessarily imply the rejection of the efficient market hypothesis. It is also important to note that factors other than nonlinear deterministic chaos can contribute to the rejection of the iid hypothesis, such as government policies and changes in the economic environment. These factors can lead to instability in the hourly data of Bitcoin pricing or linear dependence/nonlinear stochastic processes between past and future data, requiring the use of more complex models to account for the effects, such as nonlinear moving average (NMA), threshold autoregressive (TAR). To investigate these factors, the study conducts tests for stationarity, autocorrelation and linear dependence, chaos, and heteroscedasticity and nonlinear multiplicative dependence.

First, the stationarity of the hourly data of Bitcoin is tested using the Dickey-Fuller (1979) unit root test, which confirms the presence of stationarity in the sequence. To determine if stationarity is the main cause of the rejection of the iid assumption in the hourly data of Bitcoin, the data is transformed by taking the natural logarithm and then first-order differencing to obtain a stable sequence of returns. The BDS test is then performed on the transformed series and compared to the original data. Although the BDS values of the stationary series are significantly smaller than those of the original data, all of them still exceed the critical values, indicating that factors other than stationarity contribute to the rejection of the iid assumption.

Next, tests for autocorrelation and linear dependence are conducted. To distinguish between linear and nonlinear relationships, several possible ARMA (p,q) models are fitted to the stock returns, and the best model, AR(4) with zero mean, is selected based on the AIC and BIC. The BDS test is performed on the residuals of the selected model. Since the BDS test results before and after fitting the linear model are similar, it can be concluded that there is no significant linear relationship in the return series. A GARCH(1, 2) is also fitted to the residuals of the AR(4) model, as heteroskedasticity was found within the residual squared of the AR(4) model. BDS tests on the residual and volatility still indicate the series to be non iid.

However, the absence of a significant linear relationship does not imply the absence of a nonlinear relationship. Therefore, the R/S analysis is conducted to confirm the presence of nonlinear behavior in the return series of Bitcoin. The analysis results suggest that the Bitcoin trading market does not conform to the efficient market hypothesis. To further investigate the presence of chaos in the return sequence, correlation dimension analysis is performed by comparing the correlation dimensions of the fitted GARCH(1, 2) residual series, and the volatility of the fitted GARCH(1, 2). The results indicate that the correlation dimension of the residual series is low, indicating the presence of a nonlinear system. However, the lack of convergence in the correlation dimensions of the residuals contradicts the characteristic of converging dimensions in chaotic systems. Therefore, it can perhaps be concluded that the Bitcoin time series represents a nonlinear stochastic system rather than a nonlinear deterministic chaos system.

In summary, the analysis suggests that perhaps the Bitcoin price follows a nonlinear stochastic process rather than a nonlinear deterministic process, indicating that classifying it as a chaotic system may not be correct. Consequently, it is not possible to find a deterministic model to explain the behavior of Bitcoin volatility.

6. References

Aradhyula, S. V. and M. I. Holt, GARCH Time-Series Models: An Application to Retail Livestock Prices, Western Journal of Agricultural Economics, 13, 365-374, 1988.

Baillie, R. I. and R. P. DeGennaro, Stock Returns and Volatility, Journal of Financial and Quantitative Analysis, 25, 203-214, 1990.

Baumol, W. and J. Benhabib, Chaos: Significance, Mechanism, and Economic Applications, Journal of Economic Perspectives, 3, 77-105, 1989.

Blank, S. C. "Chaos" in Futures Markets? A Nonlinear Dynamical Analysis, Journal of Futures Markets, 11,711-728, 1991.

Brock, W. A., W. D. Dechert, and J. A. Scheinkman, A Test for Independence Based on the Correlation Dimension, SSRI Working Paper #8702, University of Wisconsin-Madison, 1987.

Brock, w. A. and C. L. Sayers, Is the Business Cycle Characterized by Deterministic Chaos?, Journal of Monetary Economics, 22, 71-90, 1988.

Butler, A., A Methodological Approach to Chaos: Are Economists Missing the Point?, Federal Reserve Bank of St. Louis, 72, 36-48, 1990.

Casdagli, M., Chaos and Deterministic versus Stochastic Nonlinear Modelling, Journal of the Royal Statistical Society, B54, 303-328, 1991.

Dickey, D. A. and W. A. Fuller, Distribution of the Estimators for Autoregressive Time Series with a Unit Root, Journal of the American Statistical Association, 74, 427-431, 1979.

Engle, R., Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflations, Econometrica, 50, 987-1007, 1982.

Fama, E., Efficient Capital Markets: A Review of Theory and Empirical Work, Journal of Finance, 25, 383-417, 1970.

Feigenbaum, M. J., Univeral Behavior in Nonlinear Systems, Physica, 7D, 16-39, 1983.

Grassberger, P. and I. Procaccia, Measuring the Strangeness of Strange Attractors, Physica, 9D, 189-208, 1983.

Testing For Nonlinear Properties and Chaos Phenomenon of BitCoin

57211527 Lin, Yu-Chi
Hot Papers in Finance
c.e. Prof. Coad, Alex
d.e. Prof. Arikawa, Yasuhiro d.e. Prof. Cheung, Ming Yan William