Taylor & Francis
Taylor & Francis Group

Journal of Computer Information Systems

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ucis20

Information Efficiency in the Cryptocurrency Market:The Efficient-Market Hypothesis

Ho-Jun Kang, Sang-Gun Lee & Soo-Yong Park

JOURNAL OF COMPUTER INFORMATION SYSTEMS
https://doi.org/10.1080/08874417.2021.1872046

Information Efficiency in the Cryptocurrency Market:The Efficient-Market Hypothesis

Ho-Jun Kang[a], Sang-Gun Lee[a], and Soo-Yong Park[b]

[a]Business Administration, Sogang University, Seoul City, Republic of Korea; [b]Computer Science and Engineering, Sogang University, Seoul City, Republic of Korea

ABSTRACT

This study tested the efficient-market hypothesis (EMH) to examine information efficacy in the cryptocurrency market. We conducted three random walk tests to verify the weak-form EMH and used the event study method to test the semi-strong-form EMH. The analysis results demonstrated that 54 (6.04%) of the total of 893 cryptocurrency units satisfied the weak-form EMH, and 24 (2.695%) met the semi-strong market hypothesis. Furthermore, we found that, among the cryptocurrency exchanges that were established before November 2017, large size exchanges were more likely to satisfy the weak- and semi-strong-form EMHs.

KEYWORDS

Cryptocurrency market; efficient-market hypothesis; random walk; information efficacy

1. Introduction

In October 2008, Nakamoto introduced the first blockchain model to the world through a paper titled "Bitcoin: A Peer-to-Peer Electronic Cash System".[1] Then, in January 2009, the company unveiled the Bitcoin Core program, using the concept of "blockchain" to introduce its first cryptocurrency.

Many people were enthusiastic about this revolutionary peer-to-peer (P2P) system, followed by developers and businesspeople around the world who developed several cryptocurrencies. According to CoinMarketCap, a site that collects transaction records and price trends of the cryptocurrency, there were 2,355 current ciphers who conducted 77 USD billion worth of transaction as of November 19, 2019.[2] Despite the rapid growth of cryptocurrency, governments of many countries are carefully monitoring the risk of outflow of national wealth through customs-free transactions. Companies are interested in not only crypto technology but also the underlying blockchain technology and its role as a driving force for innovations. For example, Facebook established a blockchain payment system called Libra and try to make it available around the world.[3] However, it is facing strong internal and external oppositions, such as hearings at the U.S. Senate and House of Representatives and the EU, because it threatens the position of major currencies.[3,4]

Academic research on cryptocurrency is needed because investors are still divided over whether crypto assets are financially viable investments or just speculative assets. However, there is a paucity of empirical studies on cryptocurrency despite the its growing impact on the global economy.

Therefore, this study presents an analysis of whether the cryptocurrency market has the same characteristics as the traditional stock market which will shed new insights to all the stakeholders of cryptocurrency. Furthermore, this study will examine whether the cryptocurrency market meets the efficient-market hypothesis (EMH) proposed by Fama.[5]

In regard to the analysis method, we first used the random walk test to verify the validity of the weak-form EMH and employed the case study method to verify the validity of the semi-strong-form EMH over the entire cryptocurrency market. Then, we analyzed whether the semi-strong-form EMH is satisfied by the size (large, medium, and small) of the crypto exchange and by date (before or after November 2017).

2. Literature review

2.1 Definition of blockchain

Blockchain technology is a new type of digital call system based on cryptography and a distributed P2P network architecture.[1]

Blockchain is a system that enables all participants to store and view related data, rather than storing the data on a central server of a third party, such that new blocks are connected as a chain when a new transaction occurs or is edited to an existing transaction. Therefore, the technology is referred to as decentralized rather than centralized.

Accordingly, transactions executed in blockchain cryptography technology become effective as agreed transactions by all participants. A blockchain is a collection of created blocks that are connected in a chain, and an open distributed ledger stores fixed transaction records for a certain period in each block. Each block in the blockchain links to a previously created block, and this connection extends to the first block.

The Blockchain data structure is a long container that aggregates transactions into blocks that can be considered an open, distributed ledger, and each block contains a long list of transactions that determine the block size followed by the header containing metadata. Each block functions as a distributed ledger, which consists of the transaction

CONTACT Sang-Gun Lee slee1028@sogang.ac.kr Sogang University, Seoul City, Republic of Korea.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/08874417.2021.1872046

information processed so far, enabling transactions to be validated on participants' computers.

Blockchain is designed to continue to grow as a form of a distributed database and to prevent operators' arbitrary manipulations of distributed nodes as a list of data records. One of the most representative applications of blockchain is Bitcoin, which records the transaction process of crypto as a decentralized public ledger, in which transaction records are required to be encrypted.

2.2 Cryptocurrency market

The first bitcoin Genesis block (the first block in the blockchain) was created in January 2009, and bitcoin utilization has steadily increased since its first real-world transaction happened on May 22, 2010, when it was used to trade for pizza.[2]

The Bitcoin cryptocurrency hit a new high price of 19,475 USD on December 17, 2017, due to the global craze for cryptocurrency in November 2017. However, the price has steadily fallen since then due to regulations in each country. Researchers are conducting studies on the determinants and financial possibilities of price formation of bitcoin and crypto. Nikolaos[6] used Bitcoin and Ethereum to analyze market complexity and uncertainty concerning the transaction characteristics of the crypto market.

Bouri studied how cryptocurrencies respond when the price of one crypto changes.[7] For the seven major cryptocurrencies, the higher the price, the greater the impact on other cryptocurrencies. Sigaki conducted a clustering analysis of the EMH of cryptography to examine its efficiency[8]

Zhang studied bitcoin charts' correlation with Dow Jones and concluded that the two charts were related.[9] In an EMH study, Barbiera measured whether the market was efficient with only one bitcoin and concluded that it would achieve market efficiency after 2014[10].

Recent research on cryptography usually uses one bitcoin for analysis or only up to seven crypto-currencies to derive conclusions. This study differs from the previous studies because it conducted a thorough investigation of the prices of crypto-currencies. Table 1 provides information on literature review on the cryptocurrency market.

Table 1. Literature review of cryptocurrency market.

Author	Year	Topic	Summary
Antonakakis et al.[6]	2019	Cryptocurrency market contagion: Market uncertainty, market complexity, and dynamic portfolios	A study on the market complexity and uncertainty of the transaction characteristics of the cryptocurrency market
Bouri et al.[7]	2019	Co-explosivity in the cryptocurrency market	A study on the impact of soaring coin prices on other crypto-currency markets. Price increases in the seven major crypto-currencies has an impact on other coins.
Sigaki et al.[8]	2019	Clustering patterns inefficiency and the coming-of-age of the cryptocurrency market	Clustering analysis for the EMH of crypto. 20% of crypto-currency has 20% efficiency, 43% has 60% efficiency, and 37% has 80% efficiency.
Zhang et al.[9]	2018	The inefficiency of cryptocurrency and its cross-correlation with Dow Jones Industrial Average	Examines the correlation between Dow Jones and the crypto market and found that there is a constant correlation
Bariviera.F[28]	2017	The inefficiency of Bitcoin revisited: A dynamic approach	Review of Bitcoin's EMH. Bitcoin is considered to be efficient after 2014.
Makarov and Schoar[29]	2019	Trading and arbitrage in cryptocurrency markets	Analysis of the difference in profit-taking between exchanges in the United States and other countries by comparing opportunities for large-scale and repetitive profit-taking between exchanges in the crypto market
Bentov et al.[30]	2019	Tesseract: Real-Time Cryptocurrency Exchange Using Trusted Hardware	Proposing a reliable decentralized exchange
Ji Q et al.[31]	2019	Dynamic connectedness and integration in cryptocurrency markets	Analysis of 6 Cryptocurrencies to investigate transaction connectivity between cryptocurrency
Narayanan et al.[32]	2016	Bitcoin and cryptocurrency technologies: a comprehensive introduction	Understanding Bitcoin and Cryptographic Technology
Hileman and Rauchs[33]	2017	Global cryptocurrency benchmarking study	Understanding the Global Cryptocurrency Market
Gandal and Halaburda[34]	2014	Competition in the Cryptocurrency Market	Measuring Network Effectiveness Between Cryptocurrency.

2.3 Efficient market hypothesis (EMH)

The EMH, introduced by Eugene Fama[5] was a major sensation and has become an essential factor in the development of most financial theories. He believed that the random walk behavior of stock prices was due to market efficiency.[5,10] Market participants would not be able to reap excess revenues using the current information set alone because market prices in an efficient market could only be changed by unexpected new information. So, the hypothesis is that asset prices reach equilibrium when all market participants have complete information. The definition of the EMH is as follows:

(1) A comprehensive set of information available at point θ_t= t that affects the determination of stock prices at point t.

(2) A collection of information used by the market to determine stock prices at point θ_t^M= t.θ_t^M. is a subset of θ_t, which will be equal to or smaller than θ_t.

(3) At P_jt=t, the stock price j is j = 1,2,3, … N, and N is the total number of shares present in the market.

(4) θ_t combination probability distribution of stock prices at $t + h(h > 0)$, the time that the market evaluated at t point in time based on information θ_t^M

(5) $f_m(P_{1,t+h},\ldots,P_{N,t+h}|\theta_t)$ = A "true" combination probability distribution of stock prices at $t+h(h>0)$ point of time based on information θ_t
The process of price-forming individual shares at T-point in the market is assumed to be as follows:
(6) Based on the information θ_t^M, the market first estimates the probability distribution $f_m(P_{1,t+h},\ldots,P_{N,t+h}|\theta_t^M)$ for the stock prices at t + 1 point.
(7) From the combination probability distribution of (6), the appropriate price at t-point of time for each stock is determined by the Market Equilibrium Model, which indicates a model that is established when the demand and supply of each stock are matched at t-point.
(8) The assumption that the market is efficient in the pricing process can be expressed as $\theta_t^M = \theta_t$ information used to determine the price of the securities assessed at t-time is the same as all information θ_t^M available at t-time

The information used to determine the stock price assessed at t point is the same as all information θ_t available at t point. Consequently, market efficiency refers to knowing all available information and using it correctly, while consistent excess returns cannot be achieved on any information-based transaction. Malkiel and Fama classified the efficiency of the stock market into three categories based on the type of information θ_t.[10]

The weak-form EMH claims that θ_t contains all the information that can be obtained from past market transaction data, and all the information is already fully reflected in stock price; thus, technical analysis is useless. Because past stock price data are publicly available, it costs little to obtain the necessary information. The weak-form EMH argues that if these data convey any reliable signal about future performance, all investors will have already known how to use that signal. The verification of the weak-form EMH is performed by confirming that stock price movements are independent of historical values and trends.

The semi-strong-form EMH argues that θ_t contains all publicly available information, and all the information is reflected in stock prices. In addition to past stock prices, publicly available information includes all basic information about a company's production activities, the financial status of the company based on its balance sheet, and the status of patent holdings. According to the semi-strong-form EMH, no investor can earn higher returns than the market using publicly available information.

The strong-form EMH claims that θ_t contains undisclosed information available only to the entity's insiders and that this information is reflected in the stock price. Because corporate managers can benefit from stock trading using internal

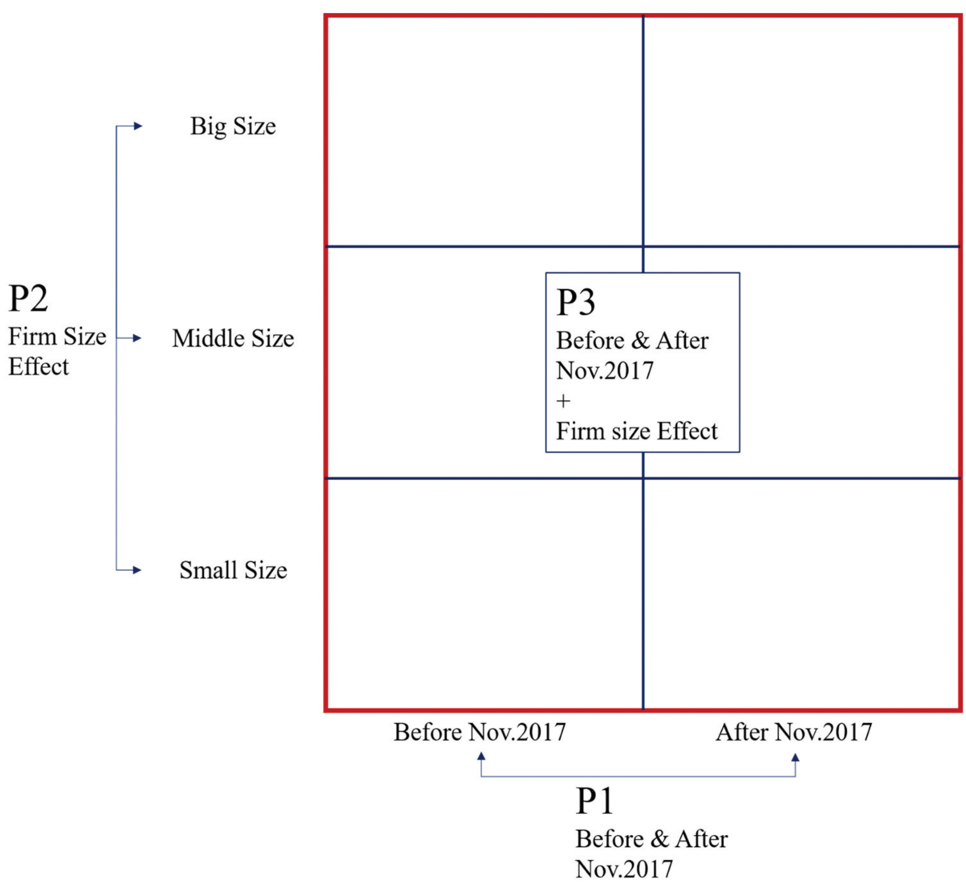

Figure 1. Research model.

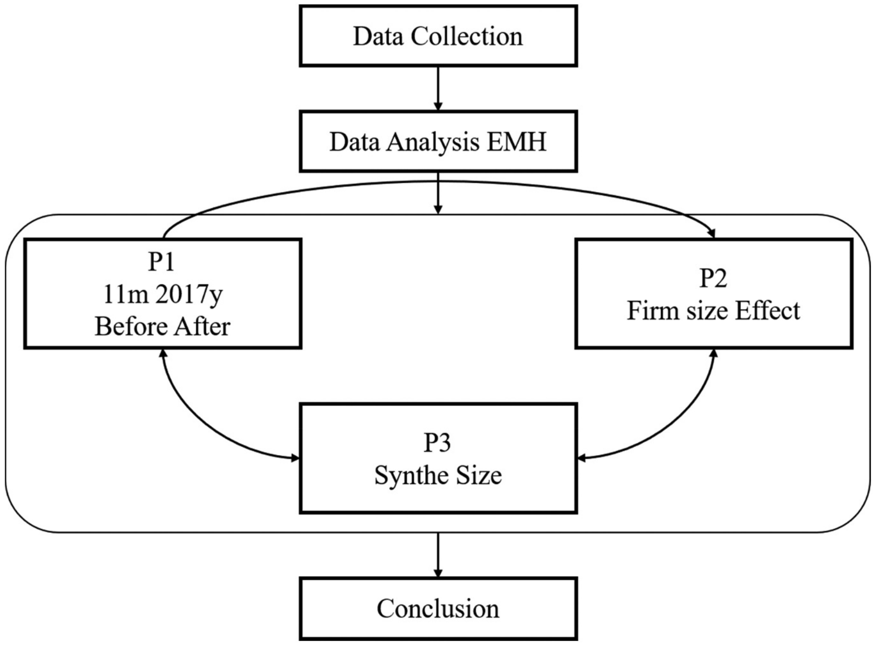

Figure 2. Research procedure.

corporate information, extreme ideal efficiency exists in reality. This situation indicates that investors will not be able to gain an excess profit through any investment strategy because the stock price always reflects fair value.

3. Research model

3.1 Research models and procedures

The research model of this study is depicted in Figure 1. We compared cryptocurrencies created before and after November 2017 to determine which group of cryptocurrencies more strongly satisfies the weak- and semi-strong-form EMHs, as claimed by Proposition 1. We also tested the weak- and semi-strong-form EMH based on the scale of the cryptocurrencies, as claimed by Proposition 2. Furthermore, we examined whether the cryptocurrencies created before November 2017 having a larger scale more closely satisfy the weak- and semi-strong hypotheses, as claimed by Proposition 3.

3.2 Data

Figure 2 illustrate the research procedure of this study. First, in order to collect necessary data for the analysis, we used Python 3.7 to crawl through the transaction records of 1,600 cryptocurrencies for the period between the effective trading days and the most recent trading days from CoinMarketCap, a website that provides trading data information from exchanges. The crawled data consists of 1.6 million columns and 1,600 types of cryptocurrency.

We then use the crawled data to verify our propositions. We use R 3.6.3, a popular statistical program, to conduct the analysis. Then, the random walk test was conducted from July 16, 2018, to July 16, 2019, to verify the weak-form EMH of cryptocurrency. As a case study to test the semi-strong-form EMH, we examine the recent U.S. Senate hearing on Facebook's Libra, held on July 16, 2019. Finally, we draw informed conclusions based on the analysis results.

3.3 Research methodology

3.3.1 Independence test for time series

The weak-form EMH assumed that changes in the formation of stock prices follow randomness.[5] Because the weak-form EMH suggested that the pattern of a stock price is already included in its present value, everyone has access to the same amount of information and can predict stock prices with efficiency unless new information is revealed.

Therefore, the pattern of a stock price is random, and the stock price will revert to zero in the long term. Consequently, it is assumed that prediction of the stock price is meaningless, and thus randomness testing is generally used to test the weak-form EMH.

The definition of the EMH, "all available information is fully reflected in the current stock prices," was interpreted as a proposition that continuous price changes and short-term returns are independent. The random walk model is used to examine whether the conditional mean of changes in a stock price depends on a change in the stock price that has been realized in the past. Therefore, we attempt to examine the weak-form efficient-market hypothesis using the following three methodologies:

A. Nonparametric independence test between time-series observations (runs test)

The runs test is a methodology proposed by Wald and Wolfowitz as a nonparametric test.[11] The null hypothesis of the runs test is as follows:

H_0: The sequence was produced in a random manner ($\tau_a < \tau < \tau_b$)

H_1: The sequence was not produced in a random manner ($\tau_a \geq \tau$ or $\tau_b \leq \tau$)

This test simplifies the data by dividing all data into only two groups, represented by 1 and 0. In this study, the calculation was performed using 1 and 0 to represent a price increase and price decrease, respectively. Thus, the modifier of the yield indicator (I_t) of each crypto is as follows:

$$I_t = \begin{cases} 1, & \text{if } r_t > 0 \\ 0, & \text{if } r_t \leq 0 \end{cases} \qquad (a-1)$$

where r_t is the daily return rate.

In the runs test, "runs" refers to the consecutive values of 1 or 0 (ex. 1111 or 0000, not 1011, 01001). The total number of runs is $N = N_+ + N_-$, and N_+ represents the number of cryptocurrencies whose daily returns increased (number of ones). N_- is the number of cryptocurrencies whose daily returns decreased (number of zeros). The mean and variance of the runs test using N are as follows.

$$\mu = \frac{2N_+ + N_-}{N} + 1 \qquad (a-2)$$

$$\tau^2 = \frac{2N_+ + N_-(2N_+ + N_- - N)}{N^2(N-1)} = \frac{(\mu-1)(\mu-2)}{N^2(N-1)} \qquad (a-3)$$

If the runs (N) are not equally distributed in the time series (if the threshold is higher or lower than the upper limit), the alternative hypothesis "the time series is not arbitrary and is not independent" is rejected. Furthermore, the null hypothesis of time-series independence can be adopted if the threshold is within the upper and lower limits. EMH must follow a random walk. Therefore, the result of adopting H0 should come out.

B. Parametric test of autocorrelation of time series (Durbin–watson test)

The Durbin–Watson test is a methodology for assessing autocorrelation in the residuals of time-series data, meaning that the error terms are not independent if there is an autocorrelation present in the residuals.[12] Autocorrelation is assumed and tested with the error term (ε_t) in the form of the primary self-recovery function, $\varepsilon\_t = \rho\varepsilon\_(t-1) + \mu\_t$.

The hypotheses are:

H_0: No first order autocorrelation. $\rho = 0$

H_1: First order correlation exists. $\rho \neq 0$

$$dw = \frac{\sum_{t=2,n}(\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1,n}(\varepsilon_t)^2} \qquad (b-1)$$

where ε_t is the residual term estimated by ordinary least squares (OLS), and the range of statistical verification using d_W statistics is as follows:

The criteria range or DW test statistics is shown in Figure 3. As illustrated in the figure, the DW test always has a value between 0 and 4. A value of 2.0 means that there is no autocorrelation detected. Values from 0 to less than 2 indicate positive autocorrelation and values from 2 to 4 indicate negative autocorrelation.

Figure 3. Residual term estimation is the lower limit of the statistic, and is the upper limit of the statistic. EMH must follow a random walk. Therefore, the result of adopting H0 should come out.

C. Variance Lo-Mackinlay test

Lo-Mackilnay proposed the verification of the variance ratio to verify the predictability of the stock market return, which has then been widely used as an indicator of market efficiency.[13]

The Lo-Mackinlay variance ratio test is built based on the random walk hypothesis (RWH), in which the variance of X_t is linear at the data interval. the variance of $X_t - X_{t-q}$ is greater by q than $X_t - X_{t-1}$. Thus, comparing the variances of the two periods enables the determination of whether the RWH is valid for the above time series.

Suppose P_t is the dollar exchange rate of the crypto at t-hour and X_t is the natural logarithm taken at P_t value.

$$[X_t = \ln P_t] \qquad (c-1)$$

Therefore, the distributed ratio value, VR(q), is defined as follows:

$$VR(q) = \frac{\sigma^2(q)}{\sigma^2(1)} \qquad (c-2)$$

$\sigma^2(q)$ represents the variance of $X_t - X_{t-q}$ by $1/q$ time and $\sigma^2(1)$ represents the variance of $X_t - X_{t-1}$.

The hypotheses are:

H_0: Variance does not follow alignment $(VR(q) = 1)$

H_1: Variance follows alignment $(VR(q) \neq 1)($

The methods for obtaining the variance of $\sigma^2(q), \sigma^2(1)$ are as follows:

$$\sigma^2(1) = \frac{1}{nq-1}\sum_{t=1}^{nq}(X_t - X_{t-1} - \hat{\mu})2 \qquad (c-3)$$

$\hat{\mu}$ is obtained as follows:

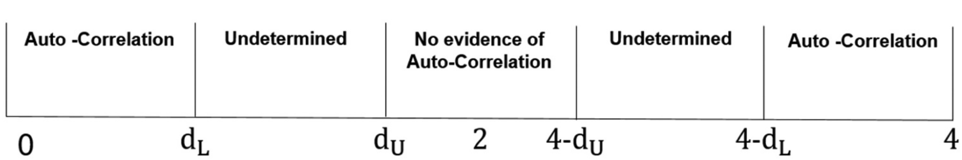

Figure 3. Durbin watson test Criteria.

$$\hat{\mu} = \frac{1}{nq}\sum_{t=1}^{nq}(X_t - X_{t-1}) = \frac{1}{nq}(X_{nq} - X_0) \qquad (c-4)$$

And $\sigma^2(q)$ is as follows:

$$\sigma^2(q) = \frac{1}{nq-1}\sum_{t=1}^{nq}(X_t - X_{t-q} - q\hat{\mu})2 \qquad (c-5)$$

Following the above steps, the value of VR(q) can be computed, and the hypothesis can be verified. EMH must follow a random walk. Therefore, the result of adopting H0 should come out.

3.3.2 Cryptocurrency 100 INDEX

The market's average daily stock market return is required before we can proceed with the event study methodology. However, because a market index of official cryptography has not yet been created, we created a new index using the Nasdaq Index Methodology used in the existing stock market. The process is as follows.[14]

First, the base period was set as August 1, 2016, when the trading of the crypto began in earnest. We assumed that the market capitalization of the top 100 of 202 cryptocurrencies was 100 at that time.

$$CINDEX_{t0} = \frac{\sum_{t=0}^{100} MarketCap}{\sum_{t=0}^{100} MarketCap} \times 100 \qquad (d-1)$$

CINDEX at the base period can be computed by Equation (1).

We then calculated the CINDEX at time N and divided it by the CINDEX at the base time.

$$CINDEX_{tn} = \frac{\sum_{t=n}^{100} MarketCap}{\sum_{t=0}^{100} MarketCap} \times 100 \qquad (d-2)$$

3.3.3 Event study

This study applied the event study methodology to verify the semi-strong-form EMH. An event study measures the information effects (the impact of the occurrence of the event on corporate value) of a particular event in a company over a short period (event period) immediately after the event occurs by the company's stock price fluctuation (stock market return).[15]

However, a certain level of stock market return is expected based on the company's risk level, even without any event, referred to as the normal rate of return. Therefore, it is necessary to estimate the information effects over the event period as an abnormal return: the gap between the normal return and the price earning ratio over the period[16]

In this study, a U.S. Senate public hearing held on July 16, 2019, was selected as the event to analyze. The normal return rate required for the computation of the abnormal return rate is calculated by the following market model regression equation:

$$\mathbf{R}^*_{it} = \alpha_i + \beta_i \mathbf{R}_{mt} + {}_{it} \qquad (e-1)$$

where R^*_{it} is the return ratio on the t-day of the cryptocurrency *I*, and R_{mt} is the rate of return on the day t market. We used the existing index of the market to create the cryptocurrency index using the indices calculation methodology and then reconstructed the market return α_i and β_i, respectively. These reconstructions refer to the estimated regression constants and regression coefficients of the cryptocurrency i based on information on the stock market return for the estimated period before the event occurred. $_{it}$ refers to the error term of company i on day t.

In the above estimation of the regression Equation 5–1, assuming the date when the event was first announced is the base period, the data of rate of return of the cryptocurrency from 30 days (−30) to 2 days (−2) before the event day was collected and applied.

Equation 5–1 applies the estimated α_i hat and β_i hat to determine the normal return rate on day t during the event period and subtracts it from the stock price earning ratio on the same day to calculate the abnormal return rate for day t based on the following equation:

$$\mathbf{AR}_{it} = \mathbf{R}_{it} - \mathbf{R}^*_{it} \qquad (e-2)$$

The abnormal return rate of a single company derived from Equation 5–2 is divided by the number of sample companies on day t in Equation 5–3 to obtain the average abnormal return (AAR) during the event window:

$$\mathbf{AAR}_{it} = \frac{1}{\mathbf{N}}\sum_{i=1}^{\mathbf{N}} \mathbf{AR}_{it} \qquad (e-3)$$

Where AAR_{it} is the average abnormal return rate on day t of the portfolio, and N is the number of cryptocurrencies collected in this study.

This study conducted a t-test to verify the level of significance of the abnormal return rate on each transaction date, and the statistics of each test were derived by Equations 5–4 and 5–5:

$$t = \frac{AR}{\sqrt{\frac{1}{60}\sum_{t=-62}^{-} \left(R_{it} - R^*_{it}\right)^2}} \qquad (e-4)$$

$$t - value = \frac{1}{N}\sum_{i=1}^{N} t \qquad (e-5)$$

4. Proposition

Opinions are divided over the economic value of cryptocurrency. Budish argued that Bitcoin is not yet a gold-like safe asset and can hardly function as a safe asset due to technical problems.[17] Davidson, Davidson stated that blockchain technology would result in a new institutional evolution, but they were not positive about cryptocurrency.[18] Li and Wang demonstrated cryptocurrency as an asset controlled by changes in economic fundamentals and market conditions in the short term.[19] Despite the various opinions in academia, cryptocurrency is not popularly used in everyday life.

Although Bitcoin is the most commonly used cryptocurrency in the world, only 19,286 stores worldwide (as of April 14, 2020) accepted it as a form of payment, which accounts for less than 15% of the 152,720 convenience stores in the United States.[20] Furthermore, in contrast to real currencies, cryptocurrencies do not offer interest rates.[1]

Therefore, cryptocurrency has different characteristics from the assets of the traditional general market, making economic evaluations of cryptocurrency remarkably difficult. This could be the reason for the cryptocurrency Black Swan event, where none of the crypto-economic experts successfully predicted the cryptocurrency bull market or identified any macroscopic signs.

According to Shiller, word of mouth causes a sort of "social infection" in people's thoughts that eventually cause bubbles in the stock market.[21] Bubbles are followed by "fantastic" rumors to eliminate doubts, which infect people and keep the bubbles alive. People believe that these social infections will continue to cause bubbles.

Based on this perspective, we can use Google search trend to gain a preview of the bubbles in the market.[22] Under social infection theory, a trend analysis of Google, the most commonly used search engine in the world, illustrates a sharp rise in search terms between November 2017 and February 2018. Furthermore, the number of cryptocurrencies increased by more than 1,000 since November 2017, and the price of cryptocurrencies are rising.

Therefore, the value of the cryptocurrencies established between November 2017 and February 2018 is speculative. Accordingly, the period between November 2017 and February 2018 is not a rational choice to be predicted by EMH, because the price and market of cryptocurrencies were caused by speculation. Moreover, the cryptocurrencies created after November 2017 are more a means of speculation than built on technological value. Proposition 1 is as follows:

(1) P1. Cryptocurrencies created before November 2017 are more likely to satisfy the EMH than those created after November 2017.
(2) P1a. Cryptocurrencies created before November 2017 are more likely to satisfy the weak-form EMH than those created after November 2017.
(3) P1b. Cryptocurrencies created before November 2017 are more likely to satisfy the semi-strong-form EMH than those created after November 2017.

In the stock market, information accessibility is unique for each company[23] and between each internal and external stakeholder.[24] This information asymmetry enables investors to make reverse choices, which can cause substantial losses for investors.

Because investors will possibly be restricted from information if information asymmetry intensifies, the liquidity of the stock market could be adversely affected, causing problems in the market. Information asymmetry varies based on the size of the company in the stock market[25]

Arbel stated that smaller companies are more likely to lack information because they are excluded from the analysis of the capital market.[26] They argued that these companies compensate for the lack of information with higher stock price returns due to the inefficiency of valuation caused by the lack of information – referred to as the small-firm effect or firm-size effect.

The larger the company, the less information asymmetry and thus the less inefficiency in valuation. Moreover, Grobys and Sapkota suggested the possibility of research that integrates financial theories of the capital market with the cryptocurrency market.[27] Consequently, if the firm-size effect is applied to the cryptocurrency, the amount of information and the degree of efficiency will vary based on the market capitalization associated with the value of the cryptocurrency in dollars.

(1) P2. The larger the scale of the cryptocurrency, the more likely that the EMH is satisfied.
(2) P2a. The larger the scale of the cryptocurrency, the more likely that the weak-form EMH is satisfied.
(3) P2b. The larger the scale of the cryptocurrency, the more likely that the semi-strong-form EMH is satisfied.

Considering both Proposition 1 and Proposition 2, the larger the scale of cryptocurrency created before November 2017, the lower the degree of information asymmetry; thus, the cryptocurrency will be more likely to satisfy the EMH. Moreover, the top 33% of cryptocurrencies created before November 2017 are the most likely to satisfy the EMH. Proposition 3 is as follows:

(1) P3. Among the cryptocurrencies created before November 2017, the group with the largest scale will be the most likely to satisfy the efficient-market assumptions.
(2) P3a. Among the cryptocurrencies created before November 2017, the group with the largest scale will be the most likely to satisfy the weak-form EMH.
(3) P3b. Among the cryptocurrencies created before November 2017, the group with the largest scale will be the most likely to satisfy the semi-strong-form EMH.

5. Results

We chose November 2018 to November 2019 as our study period to determine the random walk of the weak-form EMH. Furthermore, we divided the data ranging from January 2014 to November 2018 by month, when real transactions occurred

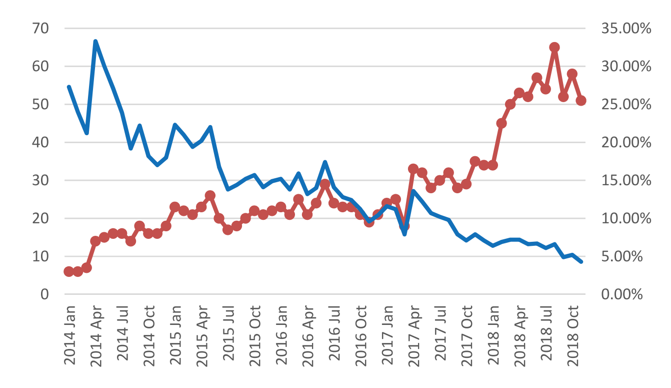

Figure 4. Trends in meeting the weak form EMH of Cryptocurrency.

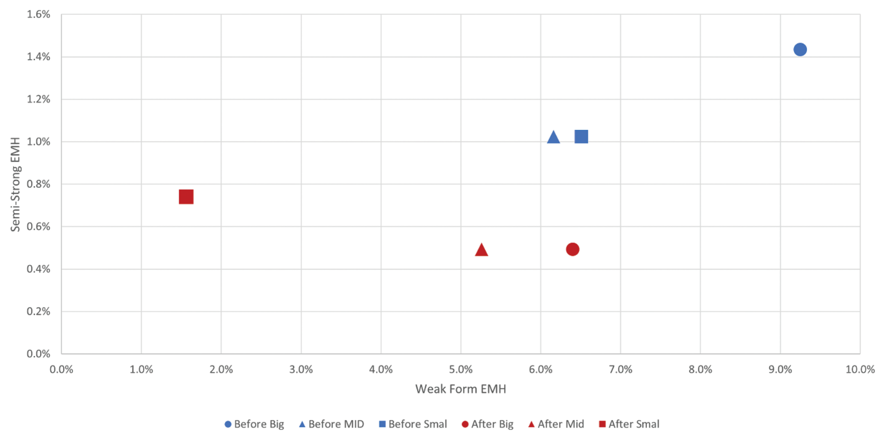

Figure 5. Summary of results.

in earnest. Eventually, we schematized the number of cryptocurrencies that passed through all the three validation tests of the weak-form EMH over the study period. We categorize the cryptocurrencies by the time period when they were created, and the number of cryptocurrencies in each category is shown in Table 2. The analyses results are as follows:

The number of cryptocurrencies that satisfy the weak-form EMH was on a steady rise since February 2014, and the efficiency plummeted sharply from July to November 2017 as the market continued to stabilize.

The results illustrate that the cryptocurrency market has been growing in the early stage due to human economic efficiency. Then, the market efficiency decreased dramatically from July to November 2017 and has since grown again.

The price fluctuation occurred because investments in the global cryptocurrency market since the end of 2017 were simple speculative investment assets rather than investment in the cryptocurrency's technological value. The results reveal that market efficiency has declined sharply since this period.

From July 2018 to July 2019, 54 of the 893 cryptocurrencies (61%) satisfied the weak-form EMH. Figure 4 illustrates that after the inefficiency of cryptocurrency reached a low in November 2017, the cryptocurrency market has been changing back to a market where information is adequately reflected.

Based on the results, the number of cryptocurrencies that satisfy the EMH steadily increased, although the percentage of cryptocurrencies that satisfy the EMH continued to decline.

The crypto market was a more economically efficient market at the initial stage and continued to be economically efficient; however, the economic efficiency decreased over time. The test results of the weak-form EMH from July 16, 2018, to July 16, 2019, were classified based on Proposition 1. Table 3 indicates the number and percentage of cryptocurrencies that satisfies the EMH within the specified time period.

The number of currencies that satisfied the weak-form EMH before November 2017 was 11% more than that after November 2017. Furthermore, for the semi-strong-form EMH verified in the event study, the number of cryptocurrencies before November 2017 was 1.76% higher. This result is due to the higher efficiency of cryptocurrencies before November 2017.

During the period from July 2018 to July 2019, only 24 of 335 cryptocurrencies (1.31%) satisfied the semi-strong-form EMH. Therefore, new information is not efficiently reflected in the cryptocurrency prices in the current cryptocurrency market.

Next, we analyzed cryptocurrencies by scale. A total of 893 cryptocurrencies were divided into three groups by scale, and an analysis was conducted on each group. The summary of the analysis results is shown in Table 4 shows the number and

Table 2. Data summary.

	Total number of cryptocurrencies	1600
	Created Before Nov. 2017	488
	Created After Nov. 2017	1112
Total number of cryptocurrencies	The test period for weak-form efficiency market hypothesis	7.2018–7.2019
Number of cryptocurrencies over the study period	Number of cryptocurrencies over the study period	893
	Created Before Nov. 2017	488
	Created After Nov. 2017	405

Table 3. Results of EMH as of november 2017.

	Before November 2017	After November 2017	Total
Number of cryptocurrencies tested	488	405	893
Number of cryptocurrencies satisfying the weak-form efficient hypothesis	31	23	54
Percentage of cryptocurrencies satisfying the weak-form efficient hypothesis	6.35%	5.65%	6.04%
Number of cryptocurrencies satisfying the semi strong-form efficient hypothesis	17	7	24
Percentage of cryptocurrencies satisfying the semi strong-form efficient hypothesis	3.48%	1.73%	2.69%

Table 4. Results of the firm-size effect of the cryptocurrency.

		Weak-form efficient hypothesis		Semi-strong efficient hypothesis	
Size	Total N	Number	Percentage	Number	Percentage
Big	298	24	8.05%	7	3.04%
Middle	298	17	5.70%	5	2.36%
Small	297	13	4.36%	5	2.70%
Total	893	54	6.04%	17	2.69%

Table 5. Weak-form hypothesis results by time and scale.

		Before November 2017 (N = 488)		After November 2017 (N = 405)		Total (N = 893)	
Size	Total N	Number	Ratio	Number	Ratio	Number	Ratio
BIG	298	16	9.25%	8	6.40%	24	8.05%
Middle	298	9	6.16%	8	5.26%	17	5.70%
Small	297	11	6.51%	2	1.56%	13	4.38%
Total	893	36	7%	18	4%	54	6%

Table 6. Semi-strong-form hypothesis results by time and scale.

		Before November 2017 (N = 488)		After November 2017 (N = 405)		Total (N = 893)	
Size	Total N	Number	Ratio	Number	Number	Ratio	Number
Big	298	7	1.43%	2	0.49%	9	3.04%
Middle	298	5	1.02%	2	0.49%	7	2.36%
Small	297	5	1.02%	3	0.74%	8	2.70%
Total	893	17	3.48%	7	1.73%	24	2.69%

percentage of cryptocurrencies that satisfies the EMH based on firm-size. The results are as follows:

The analysis results indicate that large-sized cryptocurrencies are more likely to satisfy the EMHs. Among the top 298 cryptocurrencies by scale, 8.05% satisfied the weak-form EMH, and 3.04% satisfied the semi-strong-form EMH, supporting Proposition 2. The second-highest group is the middle-sized group with 298 cryptocurrencies, in which 5.70% satisfied the weak-form EMH, and 2.36% satisfied the semi-strong-form EMH.

The larger the scale of the cryptocurrencies, the more likely they satisfy the EMH, supporting Proposition 2. Therefore, scale impacts whether the EMH is satisfied. The analysis results are classified into the weak-form and semi-strong-form EMHs, as presented in Tables 5 and 6.

First, according to Table 5, 9.25% of large-scale cryptocurrencies created before November 2017 satisfied the weak-form EMH, which is the highest among all groups.

Second, 6.51% of the 298 small scale cryptocurrencies created after November 2017 satisfied the weak-form EMH, which is the group with the second-highest percentage. The results reveal that the larger the scale of the cryptocurrency, the more likely it satisfies the weak-form EMH. Moreover, cryptocurrencies created before November 2017 are more likely to satisfy the weak-form EMH than the ones created after November 2017.

The larger the scale of the cryptocurrency and the earlier the cryptocurrency was created, the more efficient the market operates.

6. Conclusions and limitations

This study used all available cryptocurrency data to identify whether the EMHs are satisfied and whether the information is efficient. We conducted an analysis to compare (1) cryptocurrencies created before and after November 2017, and (2) cryptocurrencies of different firm sizes. Figure 5 shows the percentage of cryptocurrencies of each group that satisfied the EMH.

Proposition 1 suggested that cryptocurrencies created after November 2017 are not as efficient as those created before November 2017. The results supported Proposition 1 because the total amount of cryptocurrencies created before November 2017 are more efficient than the subsequent cryptocurrencies. Proposition 2, which claims that larger-scale cryptocurrencies are more likely to satisfy the efficient-market hypothesis, revealed that 8.04% of the large-scale cryptocurrencies satisfy the hypothesis, which is the highest among the three groups, supporting Proposition 2.

In contrast, 2.69% of the cryptocurrencies satisfy the semi-strong-form EMH. However, the market is still immature because less than 10% of the information is efficiently used, suggesting a highly-speculative market.

Consequently, investors should invest their assets with caution. Moreover, it is still challenging to apply long-term investment techniques based on efficient-market assumptions. Therefore, applying financial theories of stock market efficiency developed on the premise of the EMHs to the cryptocurrency market is not appropriate.

Therefore, the following conclusions can be drawn. First, the pass rate of the weak efficiency market hypothesis is about 6%. In the case of the current cryptocurrency market, only a few cryptocurrencies exhibit the characteristics of the weak efficiency market hypothesis. This means that the amount of information in the existing information set that can be obtained from the market transaction data is not normally reflected in the transaction price of the cryptocurrency, so it is a market where profits and losses can be obtained only with the existing information set. Second As the ratio of the semi-strong efficient market is 3.48%, information including market transaction data and publicly available information is also not normally reflected in the transaction price of cryptocurrency, so using market transaction data and publicly available information can get a loss. The contribution of this study is that it conducted a full survey of the cryptocurrency market for the first time and discovered whether the market was efficient.

With regard to limitations, although this study used data that was crawled from CoinMarketCap, the most reliable data in the cryptocurrency market, a reliability problem may exist with the data because one cryptocurrency can be traded on multiple exchanges, instead of on a single fixed exchange like the traditional market, due to the nature of cryptocurrencies. If technologies such as side chains and interchains are developed and integrated, or if decentralized exchanges appear in the future, the research could be extended.

Acknowledgments

“This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2020-2017-0-01628) supervised by the IITP(Institute for Information & communications Technology Promotion)“

References

1. Nakamoto S, Bitcoin: a peer-to-peer electronic cash system. 2008.
2. Coinmarketcap, Coinmarketcap, available at: https://coinmarketcap.com/currencies/bitcoin/, 2019(accessed).

3. Association L, Libra white paper *Internet access:https://libra.org/en-us/whitepaper/,[accessed* June 19, 2019], 2019.
4. Beresford C. Top memorable moments from facebook's libra hearings in congress. Bloomberg; 2019. https://www.bloomberg.com/news/articles/2019-07-17/top-memorable-moments-from-facebook-s-libra-hearings-in-congress.
5. Fama E, The distribution of the daily differences of the logarithms of stock prices, *PhD diss., University of Chicago.Reprinted in the Journal of Business as "The Behavior of Stock Market Prices* (38:1), 1964, 34–105.
6. Antonakakis N, Chatziantoniou I, Gabauer D. Cryptocurrency market contagion: market uncertainty, market complexity, and dynamic portfolios. Journal of International Financial Markets, Institutions and Money. 2019;61:37–51. doi:10.1016/j.intfin.2019.02.003.
7. Bouri E, Lau CKM, Lucey B, Roubaud D. Trading volume and the predictability of return and volatility in the cryptocurrency market. Finance Research Letters. 2019;29:340–46. doi:10.1016/j.frl.2018.08.015.
8. Sigaki HY, Perc M, Ribeiro HV. Clustering patterns in efficiency and the coming-of-age of the cryptocurrency market. Sci Rep. 2019;9(1):1440. doi:10.1038/s41598-018-37773-3.
9. Zhang W, Wang P, Li X, Shen D. The inefficiency of cryptocurrency and its cross-correlation with dow jones industrial average. Physica A: Statistical Mechanics and Its Applications. 2018;510:658–70. doi:10.1016/j.physa.2018.07.032.
10. Malkiel BG, Fama EF. Efficient capital markets: a review of theory and empirical work. J Finance. 1970;25(2):383–417. doi:10.1111/j.1540-6261.1970.tb00518.x.
11. Wald A, Wolfowitz J. On a test whether two samples are from the same population. The Annals of Mathematical Statistics. 1940;11(2):147–62. doi:10.1214/aoms/1177731909.
12. Durbin J, Watson GS. Testing for serial correlation in least squares regression. III. Biometrika. 1971;58:1–19.
13. Lo AW, MacKinlay AC. Stock market prices do not follow random walks: evidence from a simple specification test. Rev Financ Stud. 1988;1(1):41–66. doi:10.1093/rfs/1.1.41.
14.. NASDAQ. 1971. Nasdaq index methodology. NASDAQ.
15. Subramani M, Walden E. The impact of e-commerce announcements on the market value of firms. Information Systems Research. 2001;12(2):135–54. doi:10.1287/isre.12.2.135.9698.
16. McWilliams A, Siegel D. Event studies in management research: theoretical and empirical issues. Academy of Management Journal. 1997;40:626–57.
17. Budish E. The economic limits of Bitcoin and the blockchain. National Bureau of Economic Research; 2018. doi:10.3386/w24717.
18. Davidson S, De Filippi P, Potts J. Blockchains and the economic institutions of capitalism. Journal of Institutional Economics. 2018;14(4):639–58. doi:10.1017/S1744137417000200.
19. Li X, Wang CA. The technology and economic determinants of cryptocurrency exchange rates: the case of Bitcoin. Decis Support Syst. 2017;95:49–60. doi:10.1016/j.dss.2016.12.001.
20. Lenard J. U.S. convenience store count. NACS; 2020. https://www.convenience.org/Media/Daily/2021/Jan/14/1-US-C-Store-Counts-Stands-at-150274_Research
21. Shiller RC. Irrational exuberance. Philosophy and Public Policy Quarterly. 2000;20:18–23.
22. Preis T, Moat HS, Stanley HE. Quantifying trading behavior in financial markets using google trends. Sci Rep. 2013;3(1):1684. doi:10.1038/srep01684.
23. Yoon H, Zo H, Ciganek AP. Does XBRL adoption reduce information asymmetry? J Bus Res. 2011;64(2):157–63. doi:10.1016/j.jbusres.2010.01.008.
24. Frankel R, Li X. Characteristics of a firm's information environment and the information asymmetry between insiders and outsiders. Journal of Accounting and Economics. 2004;37(2):229–59. doi:10.1016/j.jacceco.2003.09.004.
25. Barry CB, Brown SJ. Differential information and the small firm effect. J Financ Econ. 1984;13(2):283–94. doi:10.1016/0304-405X(84)90026-6.
26. Arbel A, Carvell S, Strebel P. Giraffes, institutions and neglected firms. Financial Analysts Journal. 1983;39(3):57–63. doi:10.2469/faj.v39.n3.57.
27. Grobys K, Sapkota N. Cryptocurrencies and momentum. Econ Lett. 2019;180:6–10. doi:10.1016/j.econlet.2019.03.028.
28. Bariviera AF. The inefficiency of Bitcoin revisited: a dynamic approach. Econ Lett. 2017;161:1–4. doi:10.1016/j.econlet.2017.09.013.
29. Makarov I, Schoar A. Trading and arbitrage in cryptocurrency markets. J Financ Econ. 2020;135:293–319. doi:10.2139/ssrn.3171204.
30. Bentov I, Ji Y, Zhang F, Breidenbach L, Daian P, Juels A, Tesseract: real-time cryptocurrency exchange using trusted hardware Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, CCS '19; 2019 Nov 11–15; London, UK; p. 1521–38.
31. Ji Q, Bouri E, Lau CKM, Roubaud D. Dynamic connectedness and integration in cryptocurrency markets. International Review of Financial Analysis. 2019;63:257–72. doi:10.1016/j.irfa.2018.12.002.
32. Narayanan A, Bonneau J, Felten E, Miller A, Goldfeder S. Bitcoin and cryptocurrency technologies: A comprehensive introduction. Princeton University Press; 2016. ISBN-10:0691171696.
33. Hileman G, Rauchs M. Global cryptocurrency benchmarking study. Cambridge Centre for Alternative Finance. 2017;33:33–113.
34. Gandal N, Halaburda H, Competition in the cryptocurrency market. 2014.