Research Article

Price-Volume Relationship in Bitcoin Futures ETF Market: An Information Perspective

Xudong Wang and Xiaofeng Hui

School of Business Administration, Chongqing Three Gorges University, Chongqing, China
School of Management, Harbin Institute of Technology, Harbin, China

Correspondence should be addressed to Xudong Wang; aproton@sanxiau.edu.cn

Received 30 September 2023; Revised 14 March 2024; Accepted 18 March 2024; Published 31 March 2024

Academic Editor: Florentino Borondo

Copyright © 2024 Xudong Wang and Xiaofeng Hui. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Bitcoin futures exchange-traded funds (ETFs) are recent innovations in cryptocurrency investment. This article studies the price-volume relationship in this market from an information perspective. We first propose effective mutual information which has better estimation accuracy to analyze the contemporaneous relationship. Using half-hourly trading data of the world's largest Bitcoin futures ETF, we find that trading volume changes and returns contain information about each other and are contemporaneously dependent. Then, we employ effective transfer entropy to examine the intertemporal relationship. The results show that there exists information transfer from volume changes to returns in most of our sample period, suggesting the presence of return predictability and market inefficiency. However, information transfer in the opposite direction occurs much less frequently, and the amount is typically smaller.

1. Introduction

Exchange-traded funds (ETFs) are investment vehicles that track the performance of underlying assets [1]. Just like stocks, investors can buy or sell ETFs on stock exchanges conveniently. ETFs have become an important class of financial products. According to the Financial Times, assets invested in global ETFs exceeded $10 trillion in June 2023 [2]. The first attempt to launch a Bitcoin ETF in the U.S. market can be traced back to 10 years ago when Tyler and Cameron Winklevoss submitted their application to the U.S. Securities and Exchange Commission (SEC). Since then, multiple applications from other investment firms were also submitted. However, it was not until October 2021 that the SEC approved the first Bitcoin futures ETF, the ProShares Bitcoin Strategy ETF (BITO). It invests in Bitcoin futures traded on the Chicago Mercantile Exchange and provides a convenient and secure way to gain exposure to Bitcoin. This fund amassed $1.2 billion in assets in the first two days of trading, ranking as the second most-traded ETF launch on record [3]. Currently, due to security, manipulation, and investor protection issues, the SEC has only approved Bitcoin futures ETFs. Applications to establish spot Bitcoin ETFs in the U.S. market are still pending.

Price-volume relationship in financial markets is an important issue. It provides insights into the market structure and helps assess market efficiency [4, 5]. Previous studies on this question can be roughly grouped into two categories. The first category consists of studies that analyze the contemporaneous relationship. For example, Stosic et al. analyzed daily returns and trading volume changes in 13 stock markets and reported a cross-correlation between the two variables [6]. Zhang et al. studied the crude oil futures market and found that volume and returns were correlated [7]. Alaoui et al. examined the daily Bitcoin trading data and also detected a correlation between returns and volume changes [8]. They suggested that the two variables interacted in a nonlinear way. On the other hand, some researchers found that there was no contemporaneous dependence between price and volume in some markets [9–11].

The second category consists of studies that analyze the intertemporal relationship. For example, Hiemstra and
Jones studied the returns and volume changes in the U.S. stock market and found a bidirectional nonlinear causality between them [12]. Lee and Rui examined three stock indices in New York, Tokyo, and London markets and suggested that volume did not Granger-cause returns [13]. He et al. analyzed China’s agricultural futures markets and discovered causality between returns and volume changes in soybean, soy meal, corn, hard wheat, and strong gluten wheat futures [14]. However, they found that there was no causal relation in either direction in sugar futures. Foroutan and Lahmiri studied cryptocurrency market data and reported a causal relation from returns to volume changes of Ethereum and Litecoin during the COVID-19 period [15]. However, their results suggested that there was no causal relation between Bitcoin’s volume changes and returns.

Bitcoin futures ETFs are recent innovations. This kind of product is interesting, which is traded like stocks and involves both Bitcoin spot and futures. So what is the price-volume relationship in the Bitcoin futures ETF market? To the best of our knowledge, this question has not been studied in existing literature.

Our study aims to explore this question from an information perspective. It contributes to the literature in four ways. First, we extend the idea of Marschinski and Kantz [16] to the calculation of mutual information (MI) and propose effective mutual information (EMI). Simulation results demonstrate that EMI offers higher accuracy. Second, we use effective transfer entropy (ETE) to quantify information transfer [16, 17]. Compared with traditional Granger causality, which is only appropriate for linear relations, ETE can be used in both linear and nonlinear systems. Lungarella et al. suggested that ETE may be the first choice for complex systems with prior unknown dynamics [18]. Third, in order to avoid data discretization, which causes the loss of data information, kernel density estimation (KDE) is employed to estimate ETE and EMI. Fourth, using half-hourly trading data of the world’s largest Bitcoin futures ETF, we find that returns and volume changes contain information about each other and are contemporaneously dependent. Information predominantly flows from volume changes to returns in most of the sample period, indicating the existence of return forecastability and market inefficiency.

The rest of the paper is organized as follows. Section 2 introduces the methodology. Section 3 describes the data. Section 4 displays and discusses the empirical results. Section 5 concludes this article.

2. Methodology

2.1. Mutual Information. Let \( X = \{ x_t, t = 1, 2, \ldots \} \) be a real-valued stationary stochastic process, its uncertainty can be assessed by differential entropy (DE) as

\[
H(X) = -\int f(x_t)\log f(x_t)dx_t ,
\]

where \( f(x_t) \) is the probability density function [19].

Suppose there is another real-valued stationary process \( Y = \{ y_t, t = 1, 2, \ldots \} \), the conditional DE of \( X \) given \( Y \) is

\[
H(X | Y) = -\int f(x_t | y_t)\log f(x_t | y_t)dx_tdy_t ,
\]

where \( f(x_t | y_t) \) is the conditional probability density.

MI between \( X \) and \( Y \) can be expressed through DEs as

\[
MI(X, Y) = H(X) - H(X | Y) = H(X) + H(Y) - H(X, Y).
\]

MI \((X, Y)\) can be interpreted as the reduction of uncertainty about \( X \) which results from knowing \( Y \), indicating that it quantifies the information that the two processes share [21]. Since \( H(X) \) is larger than or equal to \( H(X | Y) \) and \( MI(X, Y) = MI(Y, X) \), MI is a nonnegative and symmetric measure. Compared with the Pearson correlation coefficient which only measures linear dependence, MI can quantify both linear and nonlinear dependence. The stronger the dependence is, the larger the value of MI [22] is.

2.2. Transfer Entropy. Let \( x^k_t = (x^1_t, x^2_{t-1}, \ldots, x^{k+1}_t) \) denote the \( k \)-lag history of \( x_t \), the entropy of \( x_{t+1} \) conditioned on \( x^k_t \) is

\[
H(x_{t+1} | x^k_t) = -\int f(x_{t+1} | x^k_t)\log f(x_{t+1} | x^k_t)dx_{t+1}dx^k_t.
\]

Thus, the transfer entropy (TE) from \( Y \) to \( X \) is [17, 19, 20]

\[
TE_{Y \rightarrow X} = H(x_{t+1} | x^k_t) - H(x_{t+1} | x^k_t, y^l_t) = H(x_{t+1} | x^k_t) - H(x_{t+1}, x^k_t, y^l_t) + H(x^k_t, y^l_t).
\]

TE\(_{Y \rightarrow X}\) can be understood as the reduction of the uncertainty of \( x^k_t \) which is derived from \( y^l_t \) that does not exist in \( x^k_t \). In this way, TE measures the predictive information transfer between two systems [23]. According to equation (5), it can be known that TE is nonnegative and nonsymmetric.

2.3. Kernel Density Estimation. Equations (3) and (5) show that MI and TE can be calculated through the sum of DEs. Due to the good performance of the KDE method, we employ it to estimate the DEs in the two equations.

For a dataset \( Z = \{ z_1, z_2, \ldots, z_n \} \) where \( z_i \in \mathbb{R}^d \), the probability function estimated by KDE with the Gaussian kernel which is widely used is

\[
\hat{f}(z) = \frac{1}{nh^d} \sum_{j=1}^{n} \frac{1}{(2\pi)^d |C|} \exp \left( -\frac{(z_i - z_j)^T S^{-1}(z_i - z_j)}{2h^2} \right),
\]

where \( C \) is the covariance matrix and \( h \) is the bandwidth which is determined by Silverman’s rule in this paper [24].
2.4. Effective Mutual Information and Effective Transfer Entropy. Due to the finite sample size, two independent variables can generate spurious nonzero MI and TE during calculation. Also, the parametric distribution of errors is unknown for MI and TE [26, 27]. To address this problem, Marschinski and Kantz introduced ETE [16]. In this paper, we take the following steps to calculate ETE [27–29]:

Step 1: Generate the surrogate series \(Y_{\text{shuffled}}\) by randomly shuffling the whole source series \(Y\) and calculate \(\text{TE}_{Y_{\text{shuffled}}} \rightarrow X\).

Step 2: Repeat Step 1 for \(M\) times. The one-sided \(p\)-value is estimated by counting the proportion of \(\text{TE}_{Y_{\text{shuffled}}} \rightarrow X\) that are greater than or equal to \(\text{TE}_{Y \rightarrow X}\).

Step 3: If the \(p\)-value is less than the chosen significance level, we reject the null hypothesis that no information is transferred from \(Y\) to \(X\) and apply equation (9) to calculate \(\text{ETE}_{Y \rightarrow X}\).

\[
\text{ETE}_{Y \rightarrow X} = \text{TE}_{Y \rightarrow X} - \frac{1}{M} \sum \text{TE}_{Y_{\text{shuffled}} \rightarrow X}. \tag{9}
\]

Otherwise, \(\text{ETE}_{Y \rightarrow X}\) equals zero. In our previous paper [30], we compared ETE with theoretical MI values and found that ETE had good accuracy.

In order to correct the bias in MI calculation, we extend the idea of Marschinski and Kantz [16] to MI and propose EMI, which is computed through similar steps. First, we generate a surrogate series \(X_{\text{shuffled}}\) by randomly shuffling \(X\) and calculate \(\text{MI}(X_{\text{shuffled}}, Y)\). Second, we repeat the previous step \(M\) times. Third, we estimate the one-sided \(p\) value through counting the proportion of \(\text{MI}(X_{\text{shuffled}}, Y)\) that are greater than or equal to \(\text{MI}(X, Y)\). If the \(p\) value is smaller than the chosen significance level, we reject the null hypothesis that \(\text{MI}(X, Y) = 0\) and use equation (10) to calculate EMI.

\[
\text{EMI}(X, Y) = \text{MI}(X, Y) - \frac{1}{M} \sum \text{MI}(X_{\text{shuffled}}, Y). \tag{10}
\]

Otherwise, the EMI is zero. Following the literature [27, 29], we set \(M = 500\).

To validate the effectiveness of our algorithm, we calculate the EMI and \(\text{MI}_{\text{original}}\) (estimated by equation (3)) from samples generated by a normal bivariate distribution. The theoretical MI of the samples is \(-0.5 \log(1 - \rho^2)\) where \(\rho\) is the correlation coefficient [31]. In our simulation, \(\rho\) is set in the range from \(-0.95\) to \(0.95\) with the increment of \(0.1\). For each \(\rho\), we generate 50 samples. The length of each sample is 200. Also, the results of average EMI, \(\text{MI}_{\text{original}}\), and theoretical MI are plotted in Figure 1. We can observe that \(\text{MI}_{\text{original}}\) is always overestimated and larger than the theoretical MI. So it is necessary to correct the bias in \(\text{MI}_{\text{original}}\).

On the other hand, EMI is more accurate and very close to the theoretical MI value, indicating that the bias is well corrected by the procedures of our algorithm. The absolute mean error of EMI is only 0.006.

3. Data

The data used in this study are the half-hourly price and trading volume of BITO, which is currently the world’s largest Bitcoin futures ETF. According to https://Forbes.com, assets under management (AUM) of BITO are much higher than the sum of other U.S. Bitcoin futures ETFs’ AUM [32]. The period of the data is from October 19, 2021, to July 14, 2023. The data are collected from the Choice database of Eastmoney Co., Ltd. which is a commercial financial information provider in China. Figures 2 and 3 display the data. It can be seen that, although similar to spot Bitcoin, the price of BITO has experienced a sharp plunge, and trading is still active. The average half-hourly trading volume of our sample is 661330.7.

In order to stationarize the data, we follow literature [12, 14, 15, 33] to calculate the logarithmic returns and volume changes:

\[
R_{i,t} = \ln P_{i,t} - \ln P_{i-1,t}, \tag{11}
\]

\[
V_{C_{i,t}} = \ln V_{i,t} - \ln V_{i-1,t}, \tag{12}
\]

where \(P_{i,t}\) and \(V_{i,t}\) are the price and trading volume of the \(i\)-th half hour of day \(t\). \(P_{0,t}\) and \(V_{0,t}\) are the data of the last half hour of the previous trading day [34–36].

Augmented Dickey–Fuller test is employed to examine the stationarity of \(R_{i,t}\) and \(V_{C_{i,t}}\) series. Schwarz information criterion is adopted to determine the lag length. The results reject the unit root null at a 1% significance level for both series, suggesting that they can be deemed stationary.

The parameters \(k\) and \(l\) in equation (5) can be determined by the lags which are required by the autocorrelation function to fall below \(1/e\) of its original value [37]. Both \(R_{i,t}\) and \(V_{C_{i,t}}\) series satisfy this requirement when the lag equals 1. Therefore, we set \(k = l = 1\), the same as the setting in the literature [38–40].

4. Empirical Results

Since the price-volume relationship is likely to be time-varying as the market evolves, a rolling window technique is employed. This technique is widely used in finance research [39, 41, 42]. Figure 4 shows its procedure [42]. Suppose there is a time series \(X = \{x_1, x_2, \ldots, x_n\}\). We use \(w\) to represent the window size which is the number of data points contained in each window. \(s\) represents the rolling step which is the number of data points that the window slides forward each time. So the \(k\)-th window starts at \(x_{1+(k-1)s}\) and ends at
Also, the \((k + 1)\)-th window starts at \(x_{w+1}\) and ends at \(x_{w+k}\). EMI and ETE between BITO’s trading volume changes and returns are calculated over each window. EMI quantifies the contemporaneous dependence through information sharing [22]. ETE measures the intertemporal causal relationship through information transfer [43]. BITO is listed on NYSE Arca whose regular trading hours are from 9:30 to 16:00 Eastern Time. According to equation (11), for each trading day, there are 13 half-hourly returns. So we set the rolling step equal to 13, making the window slide forward by one day each time. Following Lehrer et al., we set the window size to 600 [44]. Since there are about 22 trading days per month, thus the time span of each window is approximately 2 months. The results of other window sizes are also reported in the latter part of this paper.

Figure 5 displays the result of EMI. We can observe that although the value of EMI is time-varying, it is positive in all windows, indicating that BITO’s returns and volume changes contain some information about each other. It also implies that the two variables are contemporaneously dependent. This result is consistent with the prediction of the mixture distribution hypothesis, which argues that trading volume and returns jointly depend on a common underlying variable that is usually considered as the random flow of information to the market [45–47]. The two variables respond to the new information at the same time. Therefore, they contain common information and have contemporaneous dependence [14, 47]. According to the previous introduction, BITO invests in Bitcoin futures to track the performance of spot Bitcoin.
Figure 3: Half-hourly trading volume of BITO from 2021/10/19 to 2023/7/14.

Time Series X: \( x_1, x_2, \ldots, x_n \)

Window 1: \( x_1, x_2, \ldots, x_w \)

Window 2: \( x_{1+s}, x_{2+s}, \ldots, x_{w+s} \)

Window k: \( x_{1+(k-1)s}, x_{2+(k-1)s}, \ldots, x_{w+(k-1)s} \)

Figure 4: The procedure of the rolling window technique.

Figure 5: Evolution of the EMI between BITO's trading volume changes and returns at the significant level of 0.05. The horizontal axis is the start date of each rolling window.
Figure 6: Evolution of the ETE between BITO’s returns and trading volume changes at the significance level of 0.05. The horizontal axis is the start date of each rolling window.

Figure 7: Continued.
Figure 7: Evolution of the EMI (a) and ETE (b) between BITO’s returns and trading volume changes at the significance level of 0.01. The horizontal axis is the start date of each rolling window.

Figure 8: Continued.
Karpoff et al. argued that there was an absence of contemporaneous dependence between volume and returns in futures markets due to the symmetric cost of taking long and short positions [4, 10]. However, Kao et al. detected a contemporaneous return-volume relationship in the Bitcoin futures market [48]. In the Bitcoin spot market, the contemporaneous correlation between volume and return is also found [8, 49].

The results of ETE are presented in Figure 6. We can observe that the ETE from BITO’s volume changes to its returns are positive in most of the windows. It can be inferred that there exists information transfer in this direction, indicating that past volume changes can help forecast future returns. This phenomenon is consistent with the prediction of the sequential information arrival hypothesis. This theory assumes that individuals in the market receive information sequentially and randomly. As each individual receives the new information, he takes positions and adjusts his portfolios accordingly. This causes shifts in the demand and supply and a series of momentary equilibriums before the final information equilibrium occurs. Therefore, past volume is conductive for making new decisions and contains information about future returns [14, 47, 50–52]. Besides, Blume et al. theoretical model suggests that volume can provide information on information quality that cannot be derived from the price statistic [53]. Suominen’s model also shows that past trading volume can help assess the availability of private information which can be used by traders to adjust their trading strategies [54]. The two models also indicate the return predictability from volume. On the other hand, according to the efficient market hypothesis, if a market is efficient, the price should fully reflect all available information. Therefore, there should be no return predictability [55]. So we can infer that the market is not efficient in these periods. This finding also suggests that the role of trading volume should be considered when we design return forecasting algorithms.

In Figure 6, we can also observe that the ETE from BITO’s returns to its volume changes is usually smaller and zero in many windows, indicating that information transfer occurs less frequently in this direction, and the magnitude is weaker. These findings imply that the dominant direction of the information flow is from volume changes to returns. This result differs from the dominant direction found in the Bitcoin spot market by Sahoo and Sethi [5, 56]. Sahoo found that returns Granger caused volumes but not vice versa [5]. Fousekis and Tzaferi found that spillover from returns to volumes was stronger than that in the opposite direction [56].

In order to check the robustness of our findings, we have also examined the results at the significance level of 0.01 and presented them in Figure 7. The results in Figure 7(a) are similar to Figure 5. Except for several windows, EMI is still positive. In Figure 7(b), we can see that the results are similar to Figure 6. The ETE from volume changes to returns is still positive in most of the windows. Also, the ETE in the opposite direction is usually smaller. Information still predominantly flows from volume changes to returns. Our results suggest that the contemporaneous and intertemporal dependence between returns and volume changes can both exist in a market, indicating that it may be not enough to describe the price-volume dynamics in a complex market with the mixture distribution hypothesis or the sequential information arrival hypothesis alone. Literature [14, 15, 57, 58] also found this phenomenon in other markets.

Referring to the literature [59, 60], we have also examined the results of different window sizes. Figure 8 shows
Figure 9: Continued.
Figure 9: Continued.
the evolution of EMI when the window sizes are equal to 300, 900, and 1200. As there are 13 half-hourly data points per trading day, the time spans of these windows are approximately 1 month, 3 months, and 4 months, respectively. In the figure, we can observe that when the window sizes are 900 and 1200, EMI is positive at all times, even with a significance level of 0.01. When the window size is 300, although EMI fluctuates more and is zero in some windows, it is still positive for most windows. Thus, it can be concluded that even in short term, returns and trading volume changes are generally contemporaneously dependent, and this relationship is more stable in long term.

Figure 9 displays the evolution of ETE when the window sizes are equal to 300, 900, and 1200. We can observe that the ETE from volume changes to returns is still often above zero. Also, the ETE from returns to volume changes is usually smaller and even equal to zero in many windows. Thus, we can still conclude that information predominantly flows from volume changes to returns. So Figures 8 and 9 also support the conclusions that we obtained before when the window size was set to 600. Besides, we can see that when the window sizes are 900 and 1200, ETE in both directions evolves more smoothly, indicating that the information flow is more stable in long term.
5. Conclusion

The price-volume relationship has received considerable attention in the fields of finance and economy. From an information perspective, this paper investigates whether there are contemporaneous and intertemporal linkages between returns and volume changes of Bitcoin futures ETFs. To address the finite sample size effect, we extend the idea of Marschinski and Kantz [16] to MI estimation and propose an improved estimator called EMI, which is then adopted to examine the contemporaneous dependence. Using the trading data of BITO, it is found that returns and volume changes contain information about each other and are contemporaneously dependent. Furthermore, we apply the ETE method to analyze the intertemporal dependence. The results show that, in most of our sample period, there is an information flow from BITO’s volume changes to returns, indicating that past volume changes can help predict future returns. On the other hand, the predictability of returns also implies that the price does not incorporate all information, suggesting the presence of market inefficiency.

Data Availability

The data used in this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by Chongqing Three Gorges University (Grant number: 09926701).

References


