Analysis of the high-frequency volume-price relationship of Bitcoin

Fang Wan[a], Chun-Xiao Nie[a,b,*]

[a] *School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310018, China*
[b] *Collaborative Innovation Center of Statistical Data Engineering, Technology & Application, Zhejiang Gongshang University, Hangzhou, 310018, China*

Abstract

Keywords: Information flow, Transfer entropy, Cryptocurrency market, Volume-price relationship

1. Introduction

Research into the volume-price relationship provides insights into the underlying mechanisms of market efficiency and liquidity formation. By quantifying the extent to which volume influences or predicts price, analysts can develop more robust models for risk management, trading strategies, and regulatory assessments.

2. Data and Software

The primary data source used in this study is the Gemini Exchange dataset available from `https://www.cryptodatadownload.com/data/gemini/`, which provides minute-level historical data from 2017 through the most recently available period for Bitcoin trading. This dataset includes timestamped open, high, low, close (OHLC) prices, and trading volume calculated in BTC and USD. The high-frequency nature of the data enables fine-grained temporal analysis and makes it particularly suitable for exploring information flow dynamics within short time intervals.

3. Method

3.1. Transfer entropyTE

We use transfer entropy (TE) to capture non-linear and directional information flow between volume and price, which is particularly well-suited for analyz-

*Corresponding author
Email address: `niechunxiao2009@163.com; niechunxiao@zjgsu.edu.cn` (Chun-Xiao Nie)

ing high-frequency minute-level data, such as that used in this study. Formally, TE from $X = \{x_t\}$ to $Y = \{y_t\}$ is defined as Eq. 1 Schreiber (2000).

$$TE_{x\to y} = \sum p(x_{t+1}, x_t^k, y_t^l) \log \frac{p(x_{t+1}|x_t^k, y_t^l)}{p(x_{t+1}|x_t^k)} \tag{1}$$

$$x_t^{'} = \begin{cases} s_1 & x_t \leq q_{0.1} \\ s_2 & x_t \in (q_{0.1}, q_{0.5}] \\ s_3 & x_t \in (q_{0.5}, q_{0.9}] \\ s_4 & x_t > q_{0.9} \end{cases} \tag{2}$$

3.2. VAR-based TE analysis

We employ VAR to filter out the linear components in the volume-price relationship, then examine whether information flow exists between the residuals to determine if nonlinear characteristics are present in the volume-price relationship. The steps are as follows.

$$\begin{bmatrix} r_t^P \\ r_t^V \end{bmatrix} = c + \sum_{i=1}^{p} A_i \begin{bmatrix} r_{t-i}^P \\ r_{t-i}^V \end{bmatrix} + \varepsilon_t \tag{3}$$

4. Results

4.1. Volume-price relationship of high-frequency data

4.2. Analysis of nonlinear volume-price relationships

4.3. Trend analysis

5. Conclusions

CRediT authorship contribution statement

Funding Sources

The work is supported by the characteristic & preponderant discipline of key construction universities in Zhejiang province (Zhejiang Gongshang University − Statistics). Fund Number : 1020JYN6523001G-026; 1020JYN6524001G-020.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. No conflict of interest have been declared by the authors.

References

References

Schreiber, T. (2000). Measuring information transfer. *Physical review letters*, *85*, 461.

Appendix