# Dynamic of Information Flow inside the Trading Volume and Price of Bitcoin: A Perspective

Frankie Buttons

April 16, 2025

**Abstract**

This paper examines the stock volume-price dynamic in EncryptoCurrency markets. The study employs Transfer Entropy (TE) to examine the temporal information flow between trading volume and price of Bitcoin, and identifies potential non-linear Predictability of future price movements. The results also suggest that the trend in this non-linear information flow may decrease when the market is confronted with huge price volatility and other market or political events asymmetriccally.

**Keywords:** Bitcoin, Non-Linear Modeling, Transfer Entropy, Trending Tnalysis, EncryptoCurrency, Predictability

## 1 Introduction

### 1.1 Background

In the field of cryptocurrency market analysis, the relationship between trading volume and asset price and whether it is predictable has garnered increasing attention. Understanding this volume-price dynamic is essential, as trading volume often reflects market participation, investor sentiment, and potential price momentum. Unlike traditional financial markets, cryptocurrencies have a continous trading environment, and exhibit higher volatility and decentralized behaviors, which amplify the significance of volume-based indicators in forecasting price movements and identifying market anomalies.

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. This study seeks to explore these interdependencies using high-frequency data, offering a deeper empirical perspective on the temporal information flow between trading volume and price in the cryptocurrency ecosystem.

## 1.2 Methodology

### 1.2.1 Transfer Entropy

We use transfer entropy (TE) to capture non-linear and directional information flow between volume and price, which is particularly well-suited for analyzing high-frequency minute-level data, such as that used in this study. Though TE measures how much the past of one variable helps predict the future of another beyond its own past, it is rather robust in account of discontinuous data with some value missed in account of Causality Estimation. Formally, TE from $X$ to $Y$ is defined as:

$$TE_{X \to Y} = H(Y_{t+1} \mid Y_t^{(l)}) - H(Y_{t+1} \mid Y_t^{(l)}, X_t^{(k)}) \tag{1}$$

with Shannon entropy $H(Z)$ defined as:

$$H(Z) = -\sum_z p(z) \log p(z) \tag{2}$$

Because we not assume linearity or symmetry in the interactions between variables like price and trading volume, this makes it capable of capturing subtle, short-lived, and potentially asymmetric dependencies that may be missed by traditional linear models. In the context of cryptocurrency markets—where investor behaviors and market responses can shift rapidly—the non-parametric and model-free nature of TE provides a robust framework for detecting temporal information flow in a dynamic environment.

### 1.2.2 Vector Autoregressive Model

To account for linear dependencies, we use a Vector Autoregressive (VAR) model on the logged profit of price and volume. The residuals from VAR, representing non-linear components, are then used for TE analysis again, in order to identify whether the Volume-Price dynamic is predictable through linear model like VAR. And a VAR($p$) model is given by:

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

## 1.3 Data Processing

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.

To ensure that the statistical properties of the time series align with the assumptions of our transfer entropy framework, we preprocess the raw price and volume series by taking their logarithmic first differences. This transformation not only stabilizes the variance and reduces heteroscedasticity but also helps achieve stationarity—a crucial condition for many time-series analytical techniques. By working with log returns of price and log-differenced volume, we aim to capture meaningful fluctuations and directional information flow between trading activity and market valuation.

## 2 Method Implementation

Our Analysis mainly includes two parts:

1. We grouped the minute data by date and use Transfer Entropy to analyze the bidirectional information flow between the base logged profit seres of price and volume using `RTransferEntropy` package. And identify the trend in Transfer Ebtropy from the results of TE calculation using `trend` package. And apply significance analysis towards the results. In order to capture the implict trend in the price-volume dynamics, we do basic satistical analysis over the TEresults from the first part, and perform proper configured Seasonal Mann-Kendall Trend Test over those TErelated series.

2. The second part is to use VAR to filter the linear components of the base logged profit series of both price and volume, and extract the residules(the non-linear components). Then we use the residules of VAR to analyze the information flow between the residules of price and volume use `RTransferEntropy` package. We also apply significance analysis towards the results,

## 3 Results

Through the calculation above, We firstly identified that the information flow between the base logged profit series of TE from price and volume is generaly higher than the inverted one, which indecates The Causality from price to trading volume is dominant in the whole period that have been analyzed, thus

## 4 Conclusion