Digital Finance
https://doi.org/10.1007/s42521-023-00085-x

ORIGINAL ARTICLE

What drives cryptocurrency returns? A sparse statistical jump model approach

Federico P. Cortese[1] · **Petter N. Kolm**[2] · **Erik Lindström**[3]

Received: 14 January 2023 / Accepted: 26 April 2023

Abstract

We apply the statistical sparse jump model, a recently developed, interpretable and robust regime-switching model, to infer key features that drive the return dynamics of the largest cryptocurrencies. The algorithm jointly performs feature selection, parameter estimation, and state classification. Our large set of candidate features are based on cryptocurrency, sentiment and financial market-based time series that have been identified in the emerging literature to affect cryptocurrency returns, while others are new. In our empirical work, we demonstrate that a three-state model best describes the dynamics of cryptocurrency returns. The states have natural market-based interpretations as they correspond to bull, neutral, and bear market regimes, respectively. Using the data-driven feature selection methodology, we are able to determine which features are important and which ones are not. In particular, out of the set of candidate features, we show that first moments of returns, features representing trends and reversal signals, market activity and public attention are key drivers of crypto market dynamics.

Keywords Clustering · Blockchain · Cryptocurrencies · Feature selection · Regime switching · Unsupervised learning

JEL Classification G10 · C32

✉ Federico P. Cortese
federico.cortese@unimib.it

Petter N. Kolm
petter.kolm@nyu.edu

Erik Lindström
erik.lindstrom@matstat.lu.se

1 Department of Economics, Management and Statistics, University of Milano-Bicocca, Via Bicocca degli Arcimboldi, 8, 20126 Milan, Italy

2 Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA

3 Centre for Mathematical Science, Lund University, Sölvegatan 18A, 22362 Lund, Sweden

Published online: 20 May 2023

MSC Classification 91G15 · 62H30 · 62M05 · 62P05 · 68T99

1 Introduction

Since the introduction of Bitcoin in 2009, which operates with a decentralized ledger system known as the blockchain, cryptocurrencies have attracted much attention, and today some market participants argue that they constitute a separate asset class (Bianchi, 2020; Pele et al., 2021). Following the introduction of Bitcoin, many other cryptocurrencies have been implemented, collectively referred to as altcoins. By March 2022, there were approximately 18K cryptocurrencies in total.

Time series analysis of cryptocurrency dynamics show they exhibit regime switching, structural breaks and jumps in both returns and volatility (Ardia et al., 2019; Chaim & Laurini, 2018; Shen et al., 2020) (see Fig. 1).

In his seminal work, Hamilton, (1989) suggests that the dynamics of financial returns can be described by Markovian regime-switching processes, with drastic breaks associated with events like economic crises or political events. Later, Rydén et al., (1998) demonstrate empirically that a simple *hidden Markov model* (HMM) can reproduce most of the common stylized facts in asset returns (cf. Cont, 2001 and Lindström et al., 2015, Chapter 1).

Figà-Talamanca et al., (2021) adopt a multivariate approach to demonstrate the presence of common market regimes among cryptocurrencies. They analyze first differences of Bitcoin, Ethereum, Litecoin, and Monero prices, and conclude that a multivariate generalized white noise Markov switching model with three states best fits the data. Moreover, they characterize the regimes in terms of their different state-conditional volatilities. Koki et al., (2022) estimate several different HMMs for the purpose of forecasting Bitcoin, Ethereum and Ripple and demonstrate that a four-state specification provides the best out-of-sample performance. However, the statistical properties of the hidden states are not consistent across the three cryptocurrencies, making an interpretation of the latent states as different economic regimes difficult. An alternative explanation to their empirical results is that the switching dynamics is governed by some exogenous process(es) not included in their models.

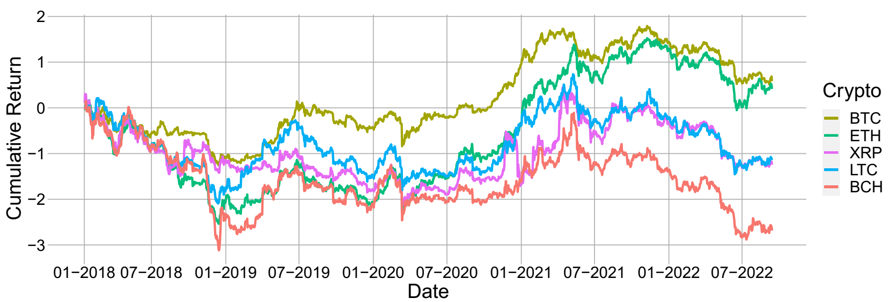

Fig. 1 Cumulative log-returns of the five cryptocurrencies Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), and Bitcoin Cash (BCH), over the period January 2018 - September 2022

While the understanding of the key drivers of cryptocurrency returns is still emerging, the literature has identified a number of crypto-, sentiment- and financial market-related features that impact cryptocurrency market dynamics (see, for example Koki et al., (2022); Yae & Tian, (2022) for an overview). Kristoufek, (2015) finds that trade and transaction volumes are positively correlated with Bitcoin returns. Similarly, Aalborg et al., (2019) suggest that frequent use of Bitcoin would increase its demand, but also that this relationship might be negative during periods of high volatility. Catania et al., (2019) and Bianchi, (2020) study the dependence of many cryptocurrencies on macroeconomic factors like S &P500 returns, market volatility (VIX) and commodities. They find a weak correlation between returns on cryptocurrencies and commodities, especially gold. Yae & Tian, (2022) demonstrate that the change in correlation between Bitcoin and S &P500 returns predicts future Bitcoin returns. In particular, they show that an increase (decrease) in correlations today suppresses (boosts) Bitcoin prices the following day. Xiong et al., (2020) provide evidence of the importance of including production cost and use value (a measure of how many people use a particular coin as a mean of exchanging value) of Bitcoin, as measured by the number of unique addresses and the Bitcoin hash rate. Several studies suggest that market sentiment plays a crucial role in determining the prices of cryptocurrencies. Kristoufek, (2013) finds a strong correlation between the number of visits to the Bitcoin Wikipedia page and price dynamics. Similarly, Aalborg et al., (2019) and Cheah et al., (2020) suggest that investor attention, as measured by Google search queries for the term "bitcoin," also impacts the cryptocurrency market. Urquhart, (2018) analyzes Google Trends search queries and suggests that large realized volatility and trading volumes increase public attention toward BTC. Likewise, Figa-Talamanca & Patacca, (2019) observe that the search volume index provided by Google is a statistically significant predictor of the conditional variance of BTC returns.

Liu & Tsyvinski, (2021) and Liu et al., (2022) find that cross-sectional factors constructed from market return, size, momentum, and public attention can forecast cryptocurrency returns. Cheah et al., (2020) and Koki et al., (2022) also find that time series momentum is a statistically significant predictor for the upward and downward trending states of crypto markets.

In this article, we aim to determine what are the most important drivers of the return dynamics of the five largest and most liquid cryptocurrencies; namely, Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), and Bitcoin Cash (BCH). For this purpose, we construct a large set of candidate features grouped into three categories: financial market, sentiment and crypto market-related features. While many of these features are drawn from the emerging literature, we also propose some that are new.

We adopt the *sparse statistical jump model* (SJM) of Nystrup et al., (2021) to select the features that best explain cryptocurrency returns. As an alternative modeling framework to HMMs, the SJMs have several advantages. First, they allow us to *simultaneously* perform parameter estimation, state-sequence decoding and feature selection. Second, Nystrup et al., (2021) show that, in contrast to many "competing" models, SJMs are more robust to misspecification and initialization, deliver acceptable performance even on smaller samples, and tend to be more efficient for high-dimensional feature vectors. Third, the estimated SJMs are easy to interpret from their conditional state dynamics and the weight associated with each feature, providing an opportunity to reconcile statistical properties and selected features with observed market behavior and economic intuition.

We demonstrate in our empirical study that when more than four hundred features are included, a three-state SJM best explains cryptocurrency returns. The three states have intuitive economic interpretations, corresponding to bull, neutral, and bear market regimes. The relevant features selected by the SJM are exponential moving averages of returns, features representing trend and reversal signals drawn from the technical analysis literature, market activity and public attention. The features that we use for representing market activity and public attention are new. Our findings are consistent with those of Figà-Talamanca et al., (2021), who also identify three common market regimes. Comparing our decoded state sequences with theirs we find that they are similar. However, in contrast to their results, we do not find the volatilities to be the key features that distinguish the market states.

The rest of the article is organized as follows. In Sect. 2, we review the mathematical formulation of SJMs, their implementation and hyperparameter tuning. In Sect. 3, we describe the candidate features we use in inferring drivers of cryptocurrency returns. We describe the preparation of our dataset and present the results from our empirical analysis in Sect. 4. Section 5 concludes. Appendix 1 provides a complete list of all the features we use in this study.

2 Methodology

Traditional regime-switching models, such as HMMs that have been used for decades in finance and economics, strike a balance between being interpretable and flexible enough to model complex non-stationary behavior. Well-known limitations of this class of models include sensitivity to model misspecification and feature selection for exogenous variables (see Zucchini et al., (2017) for a recent overview). Another issue is the difficulty to reliably estimate model parameters as the log-likelihood function is notoriously multimodal. Estimation frameworks such as direct maximization of the log-likelihood, the expectation-maximization algorithm and Bayesian approaches are compared in Rydén, (2008), with no uniform preference of any particular framework over the others.

Recently, Bemporad et al., (2018) introduced the class of *statistical jump models* (JM), nesting, e.g., HMMs. We remark that JMs are not related to jump-diffusion models, a common class of stochastic processes. Following the trend in statistics and machine learning of reformulating probabilistic models as well-behaved optimization problems, such as LASSO or support vector machines, a JM is estimated by minimizing the combined loss of the (negative) likelihood and penalties for jumping between states.

Nystrup et al., (2020) reformulate the JM as a temporal clustering model. They show in a simulation study that the JM interpretation of an HMM has many advantages, including robustness against model misspecification and poor initialization of the optimizer, as well as surprisingly rapid convergence, typically within only a few iterations. Nystrup et al., (2020) prove that the global loss can be optimized by sequentially alternating between (a) fitting the model parameters while keeping the state sequence fixed, and (b) estimating the state sequence through dynamic programming while keeping the model parameters fixed.

In a further extension, Nystrup et al., (2021) introduce the *sparse jump model* (SJM) by incorporating feature selection from the clustering literature. This is important as large scale feature selection has been infeasible for standard HMMs. They demonstrate in a series of simulation studies that SJMs outperform competing frameworks and are capable of recovering true features while rejecting false features with high probability even when considering hundreds of them.

2.1 Mathematical formulation of the SJM

The JM proposed by Bemporad et al., (2018) is governed by a latent state sequence, $\{s_t\}$, switching between K states, that each is associated with a vector of parameters, $\{\theta_k\}_{k=1}^{K}$. The model is fitted to a time series $\boldsymbol{y}_1, \dots, \boldsymbol{y}_T$, by minimizing the loss

$$\sum_{t=1}^{T-1} \left[\ell(\boldsymbol{y}_t;\, \theta_{s_t}) + \lambda \mathbb{1}_{\{s_t \neq s_{t-1}\}}\right] + \ell(\boldsymbol{y}_T;\, \theta_{s_T}), \tag{1}$$

where the hyperparameter $\lambda \geq 0$ controls the number of jumps between states and $\ell(\cdot)$ is some loss function to be specified.

Nystrup et al., (2020) consider temporal clustering by transforming the data into P standardized features, $\widetilde{\boldsymbol{y}}_{t,p} \in \mathbb{R}^p$, and using a quadratic loss function. Their corresponding objective function is given by

$$\sum_{t=1}^{T-1} \left[\|\widetilde{\boldsymbol{y}}_{t,p} - \boldsymbol{\mu}_{s_t}\|_2^2 + \lambda \mathbb{1}_{\{s_t \neq s_{t-1}\}}\right] + \|\widetilde{\boldsymbol{y}}_{T,p} - \boldsymbol{\mu}_{s_T}\|_2^2\,, \tag{2}$$

where $\boldsymbol{\mu}_{s_t}$ is the conditional mean of state s_t. It can be shown that this form of temporal clustering collapses into K-means clustering when $\lambda = 0$.

Nystrup et al., (2021) note that the *total sum of squares* (TSS) can be expressed as the sum of *within-cluster sum of squares* (WCSS) and the *between cluster sum of squares* (BCSS)

$$\sum_{t=1}^{T} \|\widetilde{\boldsymbol{y}}_{t,p} - \bar{\boldsymbol{\mu}}\|_2^2 = \sum_{t=1}^{T} \|\widetilde{\boldsymbol{y}}_{t,p} - \boldsymbol{\mu}_{s_t}\|_2^2 + \sum_{k=1}^{K} n_k \|\boldsymbol{\mu}_k - \bar{\boldsymbol{\mu}}\|_2^2\,, \tag{3}$$

where $\bar{\boldsymbol{\mu}}$ is the unconditional mean of the features and $\boldsymbol{\mu}_k$ the conditional mean of the features in the k-th state.

Building upon the work by Witten & Tibshirani, (2010) on feature selection in clustering, and using that minimizing the WCSS in equation (2) is equivalent to maximizing the BCSS (as the TSS is constant), Nystrup et al., (2021) propose to solve

$$\begin{array}{ll} \max_{\boldsymbol{\mu}_k, \{s_t\}, \boldsymbol{w}} & \boldsymbol{w}' \sum_{k=1}^{K} n_k \left(\boldsymbol{\mu}_k - \overline{\boldsymbol{\mu}}\right)^2 - \lambda \sum_{t=1}^{T-1} \mathbb{1}_{\{s_t \neq s_{t-1}\}} \\ \text{subject to} & \|\boldsymbol{w}\|^2 \leq 1, \quad \|\boldsymbol{w}\|_1 \leq \kappa \\ & w_p \geq 0 \quad \forall p \,, \end{array} \tag{4}$$

where $\boldsymbol{w}$ is the vector of feature weights, n_k denotes the number of observations belonging to the k-th cluster and the hyperparameter $1 \leq \kappa \leq \sqrt{P}$ controls the degree of sparsity of the features. Nystrup et al., (2021) show that this optimization problem can be solved by iteratively alternating between (a) fitting model parameters given the state sequence, $\{s_t\}$, and weights $\boldsymbol{w}$, (b) deriving the state sequence, $\{s_t\}$, by solving a dynamic program, essentially running the Viterbi algorithm backwards, and (c) updating the weights $\boldsymbol{w}$ using soft thresholding.

2.2 Model implementation and hyperparameters

As the SJM is unsupervised, we select its hyperparameters based on a version of the *generalized information criteria* (GIC) (Fan & Tang, 2013) for high-dimensional, penalized models suitably modified for SJMs.

Fan & Tang, (2013) consider a setup where the data $\boldsymbol{Y}$ has a number of features considerably larger than the number of observations. They define a GIC as

$$\text{GIC} = \frac{1}{T}\left\{2\left(\ell_T(\boldsymbol{Y},\boldsymbol{Y}) - \ell_T(\hat{\theta}_{\mathcal{A}},\boldsymbol{Y})\right) + a_T M\right\}, \tag{5}$$

where M is a measure of model complexity, a_T is a penalty term possibly depending on the number of observations T and features P, $\ell_T(\hat{\theta}_{\mathcal{A}},\boldsymbol{Y})$ is the log-likelihood computed using the estimated active set of features, $\mathcal{A}$, and $\ell_T(\boldsymbol{Y},\boldsymbol{Y})$ is the log-likelihood for the saturated model, which is the model obtained considering the entire set of features. Fan & Tang, (2013) note that the generalized versions of the *Akaike's information criteria* (AIC) (Akaike, 1974) or the *Bayesian information criteria* (BIC) (Schwarz, 1978) are recovered by setting M equal to the total number of parameters, and $a_T = 2$ or $a_T = \log(T)$ for AIC and BIC, respectively.

To construct a proper index, we derive expressions for the model fit and complexity and find that a BIC adjusted for SJM is remarkably efficient in selecting the correct hyperparameters values (see Cortese et al., (2023), for details), while the Fan & Tang, (2013) criteria is slightly less efficient. The AIC is designed to find the best predictive model, but does not consistently select the correct model order. This result is in agreement with the findings of Yonekura et al., (2021), who show theoretically that BIC is strongly consistent for a general class of HMMs. The modified BIC (which we denote by BIC in the following) is defined as

$$BIC = \frac{1}{T}\left\{2\left(L_T\left(\bar{\lambda},\ \bar{\kappa},\ \bar{K};\ \boldsymbol{Y}\right) - L_T(\lambda, \kappa, K; \boldsymbol{Y}) + M\log(T)\right)\right\}, \tag{6}$$

where $L_T(\lambda, \kappa, K; \boldsymbol{Y})$ denotes the estimated BCSS, $\boldsymbol{Y} \in \mathbb{R}^{T\times P}$ is the matrix of features, and λ, κ, and K are the values for the jump penalty, sparsity hyperparameter, and number of latent states, respectively. The terms $\bar{\lambda}$, $\bar{\kappa}$, and $\bar{K}$ are the SJM hyperparameters corresponding to the *saturated* model, i.e., the model obtained considering the complete set of features. While it is straightforward to set $\bar{\lambda} = 0$ and $\bar{\kappa} = \sqrt{P}$, it is not clear how to choose $\bar{K}$. Based on our experience, if the goal is to estimate a model with recurrent states, we suggest to not exceed $\bar{K} = 6$. Indeed, when $\bar{K}$ is too high, the modified BIC selects a large number of states, each one being visited only once. We define M in equation (6) by

$$M = K_0|\mathcal{A}_0| \ + \ |\mathcal{A}_0|\left(K - K_0\right) + K_0\left(|\mathcal{A}| - |\mathcal{A}_0|\right) + \sum_t \mathbb{I}_{s_t \neq s_{t-1}}, \tag{7}$$

where the three first terms come from a linear approximation of K and $|\mathcal{A}|$ near the point $(K_0, |\mathcal{A}_0|)$. This expression penalizes for increasing values of K and $|\mathcal{A}|$, the number of latent states and active features. $|\mathcal{A}|$ indirectly depends on the hyperparameter κ, and it increases with increasing values of κ. The last term in equation (7) counts the number of jumps across states and thereby depends indirectly on the jump penalty λ.

In practical applications, we suggest to select K_0 and $|\mathcal{A}_0|$ based on prior knowledge of the number of latent states and relevant features. In our empirical work, we set $|\mathcal{A}_0| = 150$ as we surmise that first- and second-order moments, correlations, volumes and momentum related features may be relevant, and these features are roughly 30 for each of the five cryptocurrencies.

The choice of K_0 is based primarily on qualitative properties. Selecting it too small will restrict the dynamics. Too large and the model does not generate persistent or even recurrent states. Rather, it segments the time series into separate blocks, thus negatively impacting model predictability. Koki et al., (2022) find that a non-homogeneous HMM with four states best describes BTC, ETH and XRP log-returns dynamics. In his comparative analysis, Bulla, (2011) observes that considering an HMM with Student-*t* conditional distributions results into selecting a fewer number of regimes. In fact, compared to Gaussian conditional distributions, model estimation is less dependent on a few extreme observations that might cause the number of states to increase. Nystrup et al., (2020) show that the JM is robust against distributional misspecification, similar to using an HMM with Student-*t* conditional distribution. Hence, we use $K_0 = 3$, reflecting our prior assumption of a modest number of states. In our setting, assuming an a priori number of states equal to 2, 3 or 4 does not significantly change the results.

3 Econometric features

We construct a large set of features as potential candidates for explaining cryptocurrency returns. Many of these features have been proposed in the literature (see Yae & Tian, (2022) for a survey), but some are new. We categorize them into three groups: financial market, sentiment, and crypto market-related features. Financial market features are based on information from the equity, fixed income, foreign exchange and commodity markets. Sentiment features proxy for investor attention toward the cryptocurrency markets. Crypto market-related features include prices and volumes of the cryptocurrencies as well as metrics related to the blockchain.

Some of our features require parameter choices before they are calculated, such as window length and the number of observations. This is addressed in our empirical study by including multiple versions of the same variable computed for different parameters, from which the feature selection algorithm then can choose the most suitable appropriate features. Next, we describe the construction of the features in each group.

3.1 Financial market features

There is an ongoing debate whether cryptocurrencies constitute a separate asset class, with some market participants arguing that due fat-tails, high kurtosis and conditional volatility of their returns (Pele et al., 2021), they behave significantly differently than traditional assets. Nevertheless, it is natural to ask whether there exists some relationship between cryptocurrencies and traditional assets classes. Therefore, we construct a number of features based on time series from the equity, fixed income, foreign exchange and commodity markets, aimed at describing cross-asset dynamics between cryptocurrencies and traditional markets. In particular, we consider the following features: first differences of WTI oil prices; first differences of 10-year minus 3-month constant maturity Treasury yields (T10Y3M); log-returns of gold; log-returns of the S &P500; log-returns of NASDAQ; log-returns of EURUSD; log-returns of JPYUSD; log-returns of CNYUSD; log-differences of VIX index; and *exponential moving averages* (EMAs) for log-differences of VIX index with half-lives $d = 1, 2, 7$, and 14 days. To proxy for possible comovements of cryptocurrencies relative to traditional asset classes (Selmi et al., 2018), we include *exponentially weighted linear and Gerber correlations*, denoted by ρ_d and g_d, (Gerber et al., 2022) of BTC log-returns with all other financial market features above with half-lives $d = 1, 2, 7$, and 14 days.

3.2 Sentiment features

Cheah et al., (2020) suggest that investor sentiment such as public attention has an impact on the cryptocurrency markets. Likewise, Aalborg et al., (2019) find that Google searches for BTC can predict BTC trading volume, while Urquhart,

(2018) and Figa-Talamanca & Patacca, (2019) provide evidence of a relationship between Google searches and volatility of BTC returns.

To proxy for public attention toward cryptocurrencies, we use the log-differences of the Google Trends indexes (GT) from the queries "bitcoin," "ethereum," "ripple," "litecoin," and "bitcoin cash." We also add a second set of public attention-related features computed as exponentially weighted linear and Gerber correlations with half-lives $d = 1, 2, 7$, and 14 days of log-differences of Google Trend indexes and log-returns of each cryptocurrency.

3.3 Crypto market-related features

This category covers features derived directly from cryptocurrency prices, trade volumes, and blockchain-related metrics. Log-differences of the total number of unique addresses with balance (AddWB) used on the blockchain aim at measuring the use value of a given coin. The number of addresses with balance is defined as the number of unique identifiers that serves as a virtual location where the coin can be sent. This metric differs from the total number of unique addresses in that it only counts wallets currently holding a particular coin, while the other one considers all addresses ever created. We use first differences of the total volume on chain (VOC), as a feature for the aggregate volume of transactions recorded on chain. Following Cong et al., (2021), we construct three value factors as the ratio between the total number of addresses and prices (AM), the number of addresses with balance and prices (UM), and the recorded volume on chain and prices (TM). We compute all the above mentioned features only for BTC, ETH, LTC and BCH due to data availability issues.

Hash rates refer to the amount of computing power used by the cryptocurrency network to process transactions and serves as a measure of the production cost of the mining process (Xiong et al., 2020). We include log-differences of hash rates (HR) for BTC and ETH.

To capture first and second moments of cryptocurrency returns, we include USD denominated daily log-returns, and EMAs of log-returns and volatilities with half-lives $d = 1, 2, 7$ and 14 days. To represent market activity, we use first differences of the logarithm of USD denominated trading volumes (V) and the corresponding EMAs with the same half-lives as above.

The results from several studies suggest that time series momentum is an important driver of crypto-returns (see, for example Liu & Tsyvinski, (2021); Liu et al., (2022); Yae & Tian, (2022)). Therefore, we include several different momentum-based features in this study. Specifically, for each of the cryptocurrencies we consider the time series momentum signal (RF) of Moskowitz et al., (2012) which is based on time series regressions with a variable lag of l. In our empirical work we use $l = 1, 2, 7$ and 14 days. Moreover, taking inspiration from the technical analysis literature, we include the relative strength index (RSI) and the moving average converge-divergence minus signal (MACDS) indicator (Wilder, 1978; Appel, 2005). The MACDS and RSI are features that represent trend and reversal signals, respectively, and are often used together to determine whether markets are in either a trending or range-bound condition. We apply the standard parameter choices when computing these features.

To proxy for illiquidity, we include the Amihud, (2002) illiquidity measure (AMIHUD), computed as the ratio of absolute daily log-returns and daily volumes for each coin.

Finally, we also construct exponentially weighted linear and Gerber correlations with half-lives $d = 1$, 2, 7 and 14 days of (a) BTC log-returns and log-returns of all other cryptocurrencies to obtain estimates of market betas, (b) log-returns and log-differences of trade volumes for each crypto, (c) BTC and ETH log-returns and log-differences of their corresponding hash rates, (d) BTC, ETH, LTC and BCH log-returns and their corresponding VOC and AddWB.

4 Empirical study

4.1 Data

The study by Alexander & Dakos, (2020) emphasizes that for empirical analysis of cryptocurrency markets the choice of data sources is critical. In particular, they advise that researchers use trade data obtained from the crypto exchanges, rather than non-trade data from coin-ranking websites and other sources where data quality is significantly lower. Barucci et al., (2022) highlight that for intraday settings, cryptocurrencies quoted against BTC, ETH or Tether (USDT) are more liquidity and therefore tend to be more accurate than those quoted against the dollar. In fact, USDT facilitates trades in cryptocurrencies as fees are lower and no bank transfers are needed.

In the present work, we take the perspective of a USD denominated investor and therefore use USD prices and volumes.[1] Following Pennoni et al., (2021), we use price and trade volume data from the Crypto Asset Lab (CAL),[2] who collect data from crypto exchanges that satisfy their reliability and liquidity criteria. We are grateful to CAL for providing us with daily volume-weighted USD denominated prices and aggregate trade volumes, recorded at midnight UTC, from the Coinbase-pro, Poloniex, Bitstamp, Gemini, Bittrex, Kraken, and Bitflyer digital exchanges.

[1] To ensure that the results of our study is not an artifact of possible differences in daily USD vs. USDT cryptocurrency quotations, we also perform our analysis with USDT denominated daily prices and volumes obtained from KuCoin. Due to data availability issues, the data for this comparative analysis covers a shorter time period, from March 7, 2019 through September 13, 2022. We observe no significant differences between the model estimated with USD or USDT denominated data. The results of this comparative analysis are available upon request.

[2] The Crypto Asset Lab is an independent lab established at the University of Milano-Bicocca; see https://cryptoassetlab.diseade.unimib.it.

Table 1 Daily unconditional means and standard deviations (SD) of the five cryptocurrency log-returns

	Mean (%)	SD (%)
BTC	0.02	4.05
ETH	0.01	5.25
XRP	0.09	6.03
LTC	0.08	5.48
BCH	0.18	6.24

Table 2 Unconditional correlation matrix of the five cryptocurrency log-returns

	BTC	ETH	XRP	LTC	BCH
BTC	1.00	–	–	–	–
ETH	0.84	1.00	–	–	–
XRP	0.65	0.70	1.00	–	–
LTC	0.81	0.84	0.71	1.00	–
BCH	0.78	0.81	0.68	0.83	1.00

Table 3 State-conditional means and standard deviations (SD) of the five cryptocurrency log-returns obtained from the SJM model

	State 1		State 2		State 3	
	Mean (%)	SD (%)	Mean (%)	SD (%)	Mean (%)	SD (%)
BTC	1.02	3.64	− 0.26	3.15	− 1.39	5.72
ETH	1.36	4.55	− 0.29	4.23	− 2.05	7.38
XRP	0.87	6.27	− 0.26	4.80	− 1.69	7.42
LTC	1.22	5.22	− 0.47	4.17	− 1.87	7.51
BCH	1.31	6.15	− 0.46	4.59	− 2.57	8.37

We obtain treasury constant maturity yields data from FRED[3]; gold and WTI oil prices, S&P500, NASDAQ and VIX levels, EURUSD, JPYUSD and CNYUSD exchange rates from Bloomberg; hash rates, number of unique addresses with balance and volumes on chain from intotheblock.com; and online search trends for the five cryptocurrencies from Google Trends.[4]

Our dataset spans the time period from January 16, 2018 through September 13, 2022, with a total number of 1,702 daily observations. As traditional financial markets are only open during business days, their corresponding time series lack values during weekends. For convenience, we impute any missing data using the `mice` **R** package (Van Buuren & Groothuis-Oudshoorn, 2011). Tables 1 and 2 provide the unconditional means, standard deviations and correlation matrix of the five cryptocurrency log-returns. From the different time series, we derive a total of 409

[3] https://fred.stlouisfed.org.

[4] https://trends.google.com.

Table 4 State-conditional correlations of the five cryptocurrency log-returns obtained from the SJM model

State 1	BTC	ETH	XRP	LTC	BCH
BTC	1.00	–	–	–	–
ETH	0.72	1.00	–	–	–
XRP	0.50	0.57	1.00	–	–
LTC	0.70	0.73	0.58	1.00	–
BCH	0.63	0.67	0.54	0.73	1.00
State 2	**BTC**	**ETH**	**XRP**	**LTC**	**BCH**
BTC	1.00	–	–	–	–
ETH	0.84	1.000	–	–	–
XRP	0.62	0.66	1.00	–	–
LTC	0.83	0.86	0.67	1.00	–
BCH	0.81	0.81	0.65	0.84	1.00
State 3	**BTC**	**ETH**	**XRP**	**LTC**	**BCH**
BTC	1.00	–	–	–	–
ETH	0.92	1.00	–	–	–
XRP	0.81	0.87	1.00	–	–
LTC	0.89	0.92	0.89	1.00	–
BCH	0.89	0.90	0.86	0.91	1.00

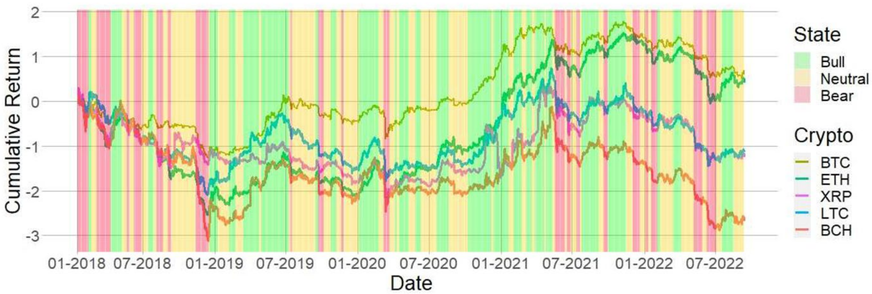

Fig. 2 Cumulative log-returns of BTC, ETH, XRP, LTC, and BCH over the period January 2018 - September 2022, together with the state sequence obtained from the SJM in green (bull), yellow (neutral) and red (bear) (color figure online)

features, all of which are stationary (see Sect. 3 for a description of the construction of the features and Appendix 1 for a complete list).

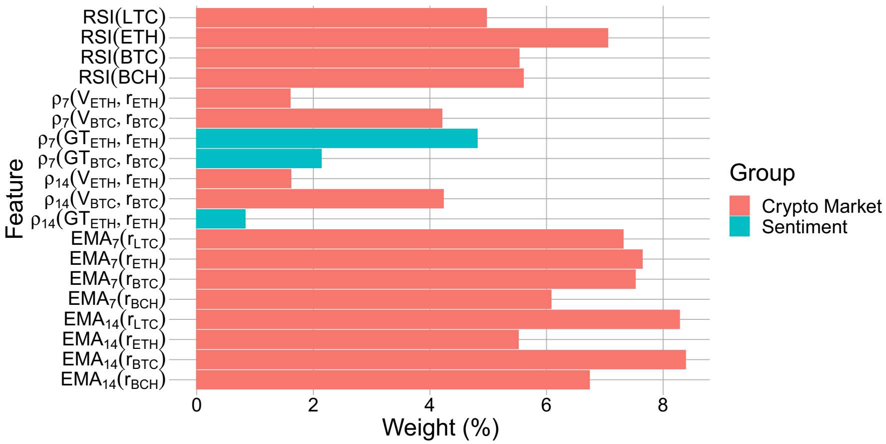

Fig. 3 Estimated weights of the relevant features in the three-state model. The selected features are RSIs for BTC, ETH, LTC and BCH; 7- and 14-day exponentially weighted linear correlations of log-differences of volumes and log-returns for BTC and ETH; 7-day exponentially weighted linear correlations of log-differences of GT and log-returns for BTC and ETH; 14-day exponentially weighted linear correlations of log-differences of GT and log-returns for BTC; 7- and 14-day EMAs of log-returns of BTC, ETH, LTC and BCH. RSI, ρ_d and EMA_d denote the relative strength index, exponentially weighted linear correlation and exponential moving average with a half-life of d days, respectively

4.2 Results

We use the `Python` implementation of the SJM from Nystrup et al. (2021), available from their online supplementary material.[5] The estimation of the SJM requires 3.1 s on a 16-core Intel i7-8750 H with 16 GB of RAM. For the remainder of this section, we discuss the results from the model.

Based on the BIC, we select the model with $K = 3$ states and with $\lambda = 5$, $\kappa = 4$. The daily state-conditional means and standard deviations of the returns are noticeably increasing from state 1 to state 3, as shown in Table 3.

In Table 4, we observe that the correlations of the cryptocurrency log-returns increase from the first through the third state, where the correlations in the third state are remarkably high. This increase in correlations is consistent with the asymmetric correlation phenomena *within* equities and other asset classes, as well as *across* asset classes (see, for example Erb et al. (1994); De Bandt and Hartmann (2000); Cappiello et al. (2006)).

Figure 2 depicts the decoded state sequence together with the cumulative log-returns of BTC, ETH, XRP, LTC and BCH.

Throughout our sample of 1702 observations, the model spends $38.9\%, 41.8\%$ and 19.3% of its time in each of the three different states, where each visit lasts for an average of 26.5, 17.8 and 16.4 days.

[5] https://www.sciencedirect.com/science/article/pii/S0957417421009647#appSB.

Table 5 Estimated weights and state-conditional values of the selected features. RSI, ρ_d and EMA_d denote the relative strength index, exponentially weighted linear correlation and exponential moving average with a half-life of d days, respectively

Feature	Weight (%)	Bull	Neutral	Bear
RSI(LTC)	5.0	63.03	43.96	31.81
RSI(ETH)	7.0	66.68	45.03	31.61
RSI(BTC)	5.5	66.46	45.71	31.58
RSI(BCH)	5.6	62.55	43.55	28.60
$\rho_7(V_{\text{ETH}}, r_{\text{ETH}})$	1.6	0.21	− 0.09	− 0.32
$\rho_7(V_{\text{BTC}}, r_{\text{BTC}})$	4.2	0.24	− 0.07	− 0.36
$\rho_7(\text{GT}_{\text{ETH}}, r_{\text{ETH}})$	4.8	0.31	− 0.08	− 0.35
$\rho_7(\text{GT}_{\text{BTC}}, r_{\text{BTC}})$	2.1	0.26	− 0.10	− 0.38
$\rho_{14}(V_{\text{ETH}}, r_{\text{ETH}})$	1.6	0.22	− 0.06	− 0.29
$\rho_{14}(V_{\text{BTC}}, r_{\text{BTC}})$	4.2	0.18	−0.04	− 0.29
$\rho_{14}(\text{GT}_{\text{ETH}}, r_{\text{ETH}})$	0.8	0.14	− 0.08	− 0.27
$\text{EMA}_7(r_{\text{ETH}})$	7.6	1.08%	− 0.25%	1.50%
$\text{EMA}_7(r_{\text{LTC}})$	7.3	0.92%	− 0.34%	− 1.56%
$\text{EMA}_7(r_{\text{BTC}})$	7.5	0.80%	− 0.15%	− 1.18%
$\text{EMA}_7(r_{\text{BCH}})$	6.1	0.98%	− 0.39%	− 2.04%
$\text{EMA}_{14}(r_{\text{ETH}})$	5.5	0.78%	− 0.17%	− 1.00%
$\text{EMA}_{14}(r_{\text{LTC}})$	8.3	0.65%	− 0.27%	− 1.12%
$\text{EMA}_{14}(r_{\text{BTC}})$	8.4	0.60%	− 0.10%	− 0.86%
$\text{EMA}_{14}(r_{\text{BCH}})$	6.7	0.65%	− 0.33%	− 1.46%

Together, the conditional statistics above suggest an interpretation of the states as distinct market regimes, where the first, second and third states represent a bull, neutral, and bear market regime, respectively. While the first regime (bull market) has positive average return and moderate volatility, the second regime (neutral market) is characterized by an average return slightly below zero under moderate volatility but with higher correlations than the first regime. In contrast, the third regime (bear market) is associated with a significant average negative return and high volatility, approximately twice the magnitude of the volatilities observed in the two other regimes. In addition, cryptocurrency returns are highly correlated in the bear market regime, suggesting that there is little to no cross-sectional diversification during bad times.

Next, we turn to examining the features selected by the SJM. From the 409 features, the model selects 19 as being relevant (see Fig. 3 for a depiction of the feature weights of the relevant features).

We observe that the 7- and 14-day EMAs of log-returns are the most relevant, with RSIs following thereafter. The state-conditional values of the relevant features are given in Table 5.

These values are consistent with the bull (positive trend), neutral (range-bound) and bear (negative trend) regime interpretations above. The state-conditional 7- and 14-day EMAs for log-returns are similar in sign and magnitude to the state-conditional means and are consistent with the upward and downward momentum observed in the first and third regime.

The state-conditional RSI for BTC are 66.46, 45.03 and 32.58, consistent with the bull (positive trend), neutral (range-bound) and bear (negative trend) regime interpretations above. State-conditional RSIs for the other cryptocurrencies have similar magnitudes.

We observe that the 7-day exponentially weighted correlation of log-differences of Google Trend index and log-returns of BTC is relevant. The state-conditional values are equal to 0.26, − 0.10 and − 0.38 in the bull, neutral and bear market states, respectively. This suggests that public attention affects the evolution of crypto markets, especially during downward and upward market trends. The model selects the 7- and 14-day exponentially weighted correlation of log-differences of Google Trend index and log-returns of ETH, having state-conditional values similar to that of $\rho_7(\mathrm{GT}_{\mathrm{BTC}},\ r_{\mathrm{BTC}})$. Similarly, the 7- and 14-day exponentially weighted linear correlations of log-returns and log-differences of the BTC and ETH trade volumes are also selected. The state-conditional values in the bull, neutral and bear market states of the 7-day correlations for BTC are 0.24, − 0.07 and − 0.36, respectively. The corresponding values for the 14-day correlations of BTC and for the 7- and 14-day correlations of ETH are similar in magnitude. Finally, we note the model does not select any features from the group of financial market features (cf. Sect. 3.1). In fact, the average state-conditional correlations of BTC log-returns and each of the financial features are close to zero in all the three regimes.

4.3 Discussion

The SJM model distinguishes three distinct regimes driven by upward, downward and sideways trends, suggesting that time series momentum is a key driver of cryptocurrencies. Notably, the presence of time series momentum is well-established in traditional asset classes such as equity, currency, commodity, and fixed income markets (Moskowitz et al., 2012; Babu et al., 2020). More recently it has also been shown to be prevalent in the crypto markets (Cheah et al., 2020; Liu & Tsyvinski, 2021; Liu et al., 2022; Koki et al., 2022). Theories of sentiment in the behavioral finance literature suggests time series momentum may result from initial under-reaction followed by delayed over-reaction,[6] Below we provide some support for this explanation and discuss how the interplay between institutional investors, who predominately act as market makers in these markets, and retail investors, who are holding positions longer, contribute to the trends.

While the interest of larger financial institutions in the crypto markets is growing, Karniol-Tambour et al., (2022) estimate that as of January 2022 only around 5% of Bitcoin is *held* by institutional investors. However, although institutions do

[6] Under-reaction is the failure of markets to fully react to new information and can occur through several behavioral channels, including gradual dissemination of news (Hong & Stein, 1999) adherence to prior beliefs and cognitive biases such as anchoring (Barberis et al., 1998), or selling profitable assets too early while holding on to losing ones too long (Shefrin & Statman, 1985). Conversely, over-reaction is the tendency of markets reacting too strongly to new information and can result from excessive optimism and self-attribution biases (Daniel et al., 1998), herding behavior (Bikhchandani et al., 1992), positive feedback trading (De Long et al., 1990), or market sentiment (Baker & Wurgler, 2006).

not appear to have the largest share of holdings in the crypto markets, in the last few years they are generally believed to be the dominant players when it comes to trading volumes (e.g., market making). In 2021, institutions traded over one trillion dollars worth of cryptocurrencies on Coinbase, an increase from the 120 billion dollar trading volume the previous year, and more than twice the amount traded by retail investors (half a trillion dollars) (Vigna, 2022). As far as positive momentum, Auer et al., (2022) show that an increase in the price of Bitcoin causes a significant entry of new retail investors in the crypto markets who in turn drive up prices further, consistent with a positive feedback trading explanation (De Long et al., 1990). Using a dataset on client transactions and account balances of retail customers at a large German online bank, Hackethal et al., (2022) suggest that customers investing in cryptocurrencies and cryptocurrency structured retail products are likely to exhibit investing biases consistent with naive trend-chasing and overtrading behavior (Barber & Odean, 2008). In addition, Kogan et al., (2022) show that many retail investors actually end up following momentum strategies, whether they are aware of it or not, when investing in cryptocurrencies. The authors argue that this behavior is predominantly driven by retail investors holding on to their positions, even in periods of large price moves. In particular, they do not rebalance after prices increase or double up when prices decrease.

The features representing the interaction of cryptocurrency returns with public attention and trade volumes are also selected by our model. Their state-conditional values are positive in the bull regime, negative in the bear regime, and close to zero in the neutral regime, consistent with, for example, Bianchi & Dickerson, (2019) and Smales, (2022).

Inspecting which groups of features are not selected by the SJM provides additional insight into the workings of crypto markets. Most importantly, features representing traditional asset classes are not helpful in explaining cryptocurrencies (see also Baur et al., (2018); Bianchi, (2020)) and neither is cryptocurrency return volatility-based features. That the latter are not selected is perhaps a bit surprising, especially as volatility-based features are some of the most important features for identifying regimes in the equity markets (Nystrup et al., 2020, 2021). A possible reason for their non-inclusion is that their state-conditional values are about the same in the bull and neutral regimes, with each being about half of the corresponding values in the bear regime.

5 Conclusions

We employed the statistical sparse jump model to infer key features that drive the return dynamics of the largest cryptocurrencies. Our results suggest that a model with three states provides an intuitive interpretation of these markets corresponding to bull, neutral and bear market regimes. We found that first moments of returns (but not second moments), features representing trends and reversal signals drawn from the technical analysis literature, market activity and public attention have the strongest descriptive power. The features that we use for representing market activity and public attention are new and aid in explaining cryptocurrency returns in upward and downward market trends.

These findings have practical implications for trading and risk management in the crypto market. In particular, practitioners can use the identified features to distinguish upward and downward market trends, and detect when the market switches between different regimes.

Appendix A: Feature set

Tag	Variable(s)	Transformation	Group
r_{BTC}	BTC log-ret	Log-difference	Crypto Market-Related
r_{ETH}	ETH log-ret	Log-difference	Crypto Market-Related
r_{XRP}	XRP log-ret	Log-difference	Crypto Market-Related
r_{LTC}	LTC log-ret	Log-difference	Crypto Market-Related
r_{BCH}	BCH log-ret	Log-difference	Crypto Market-Related
V_{BTC}	BTC Volume	Log-difference	Crypto Market-Related
V_{ETH}	ETH Volume	Log-difference	Crypto Market-Related
V_{XRP}	XRP Volume	Log-difference	Crypto Market-Related
V_{LTC}	LTC Volume	Log-difference	Crypto Market-Related
V_{BCH}	BCH Volume	Log-difference	Crypto Market-Related
$\mathrm{EMA}_1(r_{\mathrm{BTC}})$	BTC log-ret	1-day EMA	Crypto Market-Related
$\mathrm{EMA}_1(r_{\mathrm{ETH}})$	ETH log-ret	1-day EMA	Crypto Market-Related
$\mathrm{EMA}_1(r_{\mathrm{XRP}})$	XRP log-ret	1-day EMA	Crypto Market-Related
$\mathrm{EMA}_1(r_{\mathrm{LTC}})$	LTC log-ret	1-day EMA	Crypto Market-Related
$\mathrm{EMA}_1(r_{\mathrm{BCH}})$	BCH log-ret	1-day EMA	Crypto Market-Related
$\mathrm{EMA}_1(\sigma_{\mathrm{BTC}})$	BTC log-ret	1-day EMA volatility	Crypto Market-Related
$\mathrm{EMA}_1(\sigma_{\mathrm{ETH}})$	ETH log-ret	1-day EMA volatility	Crypto Market-Related
$\mathrm{EMA}_1(\sigma_{\mathrm{XRP}})$	XRP log-ret	1-day EMA volatility	Crypto Market-Related
$\mathrm{EMA}_1(\sigma_{\mathrm{LTC}})$	LTC log-ret	1-day EMA volatility	Crypto Market-Related
$\mathrm{EMA}_1(\sigma_{\mathrm{BCH}})$	BCH log-ret	1-day EMA volatility	Crypto Market-Related
$\mathrm{EMA}_2(r_{\mathrm{BTC}})$	BTC log-ret	2-day EMA	Crypto Market-Related
$\mathrm{EMA}_2(r_{\mathrm{ETH}})$	ETH log-ret	2-day EMA	Crypto Market-Related
$\mathrm{EMA}_2(r_{\mathrm{XRP}})$	XRP log-ret	2-day EMA	Crypto Market-Related
$\mathrm{EMA}_2(r_{\mathrm{LTC}})$	LTC log-ret	2-day EMA	Crypto Market-Related
$\mathrm{EMA}_2(r_{\mathrm{BCH}})$	BCH log-ret	2-day EMA	Crypto Market-Related
$\mathrm{EMA}_2(\sigma_{\mathrm{BTC}})$	BTC log-ret	2-day EMA volatility	Crypto Market-Related
$\mathrm{EMA}_2(\sigma_{\mathrm{ETH}})$	ETH log-ret	2-day EMA volatility	Crypto Market-Related
$\mathrm{EMA}_2(\sigma_{\mathrm{XRP}})$	XRP log-ret	2-day EMA volatility	Crypto Market-Related
$\mathrm{EMA}_2(\sigma_{\mathrm{LTC}})$	LTC log-ret	2-day EMA volatility	Crypto Market-Related
$\mathrm{EMA}_2(\sigma_{\mathrm{BCH}})$	BCH log-ret	2-day EMA volatility	Crypto Market-Related
$\mathrm{EMA}_7(r_{\mathrm{BTC}})$	BTC log-ret	7-day EMA	Crypto Market-Related

Tag	Variable(s)	Transformation	Group
$EMA_7(r_{ETH})$	ETH log-ret	7-day EMA	Crypto Market-Related
$EMA_7(r_{XRP})$	XRP log-ret	7-day EMA	Crypto Market-Related
$EMA_7(r_{LTC})$	LTC log-ret	7-day EMA	Crypto Market-Related
$EMA_7(r_{BCH})$	BCH log-ret	7-day EMA	Crypto Market-Related
$EMA_7(\sigma_{BTC})$	BTC log-ret	7-day EMA volatility	Crypto Market-Related
$EMA_7(\sigma_{ETH})$	ETH log-ret	7-day EMA volatility	Crypto Market-Related
$EMA_7(\sigma_{XRP})$	XRP log-ret	7-day EMA volatility	Crypto Market-Related
$EMA_7(\sigma_{LTC})$	LTC log-ret	7-day EMA volatility	Crypto Market-Related
$EMA_7(\sigma_{BCH})$	BCH log-ret	7-day EMA volatility	Crypto Market-Related
$EMA_{14}(r_{BTC})$	BTC log-ret	14-day EMA	Crypto Market-Related
$EMA_{14}(r_{ETH})$	ETH log-ret	14-day EMA	Crypto Market-Related
$EMA_{14}(r_{XRP})$	XRP log-ret	14-day EMA	Crypto Market-Related
$EMA_{14}(r_{LTC})$	LTC log-ret	14-day EMA	Crypto Market-Related
$EMA_{14}(r_{BCH})$	BCH log-ret	14-day EMA	Crypto Market-Related
$EMA_{14}(\sigma_{BTC})$	BTC log-ret	14-day EMA volatility	Crypto Market-Related
$EMA_{14}(\sigma_{ETH})$	ETH log-ret	14-day EMA volatility	Crypto Market-Related
$EMA_{14}(\sigma_{XRP})$	XRP log-ret	14-day EMA volatility	Crypto Market-Related
$EMA_{14}(\sigma_{LTC})$	LTC log-ret	14-day EMA volatility	Crypto Market-Related
$EMA_{14}(\sigma_{BCH})$	BCH log-ret	14-day EMA volatility	Crypto Market-Related
$EMA_1(V_{BTC})$	BTC Volume	1-day EMA Volume	Crypto Market-Related
$EMA_1(V_{ETH})$	ETH Volume	1-day EMA Volume	Crypto Market-Related
$EMA_1(V_{XRP})$	XRP Volume	1-day EMA Volume	Crypto Market-Related
$EMA_1(V_{LTC})$	LTC Volume	1-day EMA Volume	Crypto Market-Related
$EMA_1(V_{BCH})$	BCH Volume	1-day EMA Volume	Crypto Market-Related
$EMA_2(V_{BTC})$	BTC Volume	2-day EMA Volume	Crypto Market-Related
$EMA_2(V_{ETH})$	ETH Volume	2-day EMA Volume	Crypto Market-Related
$EMA_2(V_{XRP})$	XRP Volume	2-day EMA Volume	Crypto Market-Related
$EMA_2(V_{LTC})$	LTC Volume	2-day EMA Volume	Crypto Market-Related
$EMA_2(V_{BCH})$	BCH Volume	2-day EMA Volume	Crypto Market-Related
$EMA_7(V_{BTC})$	BTC Volume	7-day EMA Volume	Crypto Market-Related
$EMA_7(V_{ETH})$	ETH Volume	7-day EMA Volume	Crypto Market-Related
$EMA_7(V_{XRP})$	XRP Volume	7-day EMA Volume	Crypto Market-Related
$EMA_7(V_{LTC})$	LTC Volume	7-day EMA Volume	Crypto Market-Related
$EMA_7(V_{BCH})$	BCH Volume	7-day EMA Volume	Crypto Market-Related
$EMA_{14}(V_{BTC})$	BTC Volume	14-day EMA Volume	Crypto Market-Related
$EMA_{14}(V_{ETH})$	ETH Volume	14-day EMA Volume	Crypto Market-Related
$EMA_{14}(V_{XRP})$	XRP Volume	14-day EMA Volume	Crypto Market-Related
$EMA_{14}(V_{LTC})$	LTC Volume	14-day EMA Volume	Crypto Market-Related
$EMA_{14}(V_{BCH})$	BCH Volume	14-day EMA Volume	Crypto Market-Related
$\rho_1(V_{BTC}, r_{BTC})$	BTC log-ret BTC Volume	1-day EMA linear correlation	Crypto Market-Related
$\rho_1(V_{ETH}, r_{ETH})$	ETH log-ret ET Volume	1-day EMA linear correlation	Crypto Market-Related

Tag	Variable(s)	Transformation	Group
$\rho_1(V_{XRP}, r_{XRP})$	XRP log-ret XRP Volume	1-day EMA linear correlation	Crypto Market-Related
$\rho_1(V_{LTC}, r_{LTC})$	LTC log-ret LTC Volume	1-day EMA linear correlation	Crypto Market-Related
$\rho_1(V_{BCH}, r_{BCH})$	BCH log-ret BCH Volume	1-day EMA linear correlation	Crypto Market-Related
$\rho_2(V_{BTC}, r_{BTC})$	BTC log-ret BTC Volume	2-day EMA linear correlation	Crypto Market-Related
$\rho_2(V_{ETH}, r_{ETH})$	ETH log-ret ET Volume	2-day EMA linear correlation	Crypto Market-Related
$\rho_2(V_{XRP}, r_{XRP})$	XRP log-ret XRP Volume	2-day EMA linear correlation	Crypto Market-Related
$\rho_2(V_{LTC}, r_{LTC})$	LTC log-ret LTC Volume	2-day EMA linear correlation	Crypto Market-Related
$\rho_2(V_{BCH}, r_{BCH})$	BCH log-ret BCH Volume	2-day EMA linear correlation	Crypto Market-Related
$\rho_7(V_{BTC}, r_{BTC})$	BTC log-ret BTC Volume	7-day EMA linear correlation	Crypto Market-Related
$\rho_7(V_{ETH}, r_{ETH})$	ETH log-ret ET Volume	7-day EMA linear correlation	Crypto Market-Related
$\rho_7(V_{XRP}, r_{XRP})$	XRP log-ret XRP Volume	7-day EMA linear correlation	Crypto Market-Related
$\rho_7(V_{LTC}, r_{LTC})$	LTC log-ret LTC Volume	7-day EMA linear correlation	Crypto Market-Related
$\rho_7(V_{BCH}, r_{BCH})$	BCH log-ret BCH Volume	7-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(V_{BTC}, r_{BTC})$	BTC log-ret BTC Volume	14-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(V_{ETH}, r_{ETH})$	ETH log-ret ET Volume	14-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(V_{XRP}, r_{XRP})$	XRP log-ret XRP Volume	14-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(V_{LTC}, r_{LTC})$	LTC log-ret LTC Volume	14-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(V_{BCH}, r_{BCH})$	BCH log-ret BCH Volume	14-day EMA linear correlation	Crypto Market-Related
$g_1(V_{BTC}, r_{BTC})$	BTC log-ret BTC Volume	1-day EMA Gerber correlation	Crypto Market-Related
$g_1(V_{ETH}, r_{ETH})$	ETH log-ret ET Volume	1-day EMA Gerber correlation	Crypto Market-Related
$g_1(V_{XRP}, r_{XRP})$	XRP log-ret XRP Volume	1-day EMA Gerber correlation	Crypto Market-Related
$g_1(V_{LTC}, r_{LTC})$	LTC log-ret LTC Volume	1-day EMA Gerber correlation	Crypto Market-Related
$g_1(V_{BCH}, r_{BCH})$	BCH log-ret BCH Volume	1-day EMA Gerber correlation	Crypto Market-Related
$g_2(V_{BTC}, r_{BTC})$	BTC log-ret BTC Volume	2-day EMA Gerber correlation	Crypto Market-Related

Tag	Variable(s)	Transformation	Group
$g_2(V_{ETH},r_{ETH})$	ETH log-ret ET Volume	2-day EMA Gerber correlation	Crypto Market-Related
$g_2(V_{XRP},r_{XRP})$	XRP log-ret XRP Volume	2-day EMA Gerber correlation	Crypto Market-Related
$g_2(V_{LTC},r_{LTC})$	LTC log-ret LTC Volume	2-day EMA Gerber correlation	Crypto Market-Related
$g_2(V_{BCH},r_{BCH})$	BCH log-ret BCH Volume	2-day EMA Gerber correlation	Crypto Market-Related
$g_7(V_{BTC},r_{BTC})$	BTC log-ret BTC Volume	7-day EMA Gerber correlation	Crypto Market-Related
$g_7(V_{ETH},r_{ETH})$	ETH log-ret ET Volume	7-day EMA Gerber correlation	Crypto Market-Related
$g_7(V_{XRP},r_{XRP})$	XRP log-ret XRP Volume	7-day EMA Gerber correlation	Crypto Market-Related
$g_7(V_{LTC},r_{LTC})$	LTC log-ret LTC Volume	7-day EMA Gerber correlation	Crypto Market-Related
$g_7(V_{BCH},r_{BCH})$	BCH log-ret BCH Volume	7-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(V_{BTC},r_{BTC})$	BTC log-ret BTC Volume	14-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(V_{ETH},r_{ETH})$	ETH log-ret ET Volume	14-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(V_{XRP},r_{XRP})$	XRP log-ret XRP Volume	14-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(V_{LTC},r_{LTC})$	LTC log-ret LTC Volume	14-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(V_{BCH},r_{BCH})$	BCH log-ret BCH Volume	14-day EMA Gerber correlation	Crypto Market-Related
VOC_{BTC}	BTC on chain volume	First difference	Crypto Market-Related
VOC_{ETH}	ETH on chain volume	First difference	Crypto Market-Related
VOC_{LTC}	LTC on chain volume	First difference	Crypto Market-Related
VOC_{BCH}	BCH on chain volume	First difference	Crypto Market-Related
HR_{BTC}	BTC hash rate	Log-difference	Crypto Market-Related
HR_{ETH}	ETH hash rate	Log-difference	Crypto Market-Related
$AddWB_{BTC}$	BTC number of total addresses with balance	Log-difference	Crypto Market-Related
$AddWB_{ETH}$	ETH number of total addresses with balance	Log-difference	Crypto Market-Related
$AddWB_{LTC}$	LTC number of total addresses with balance	Log-difference	Crypto Market-Related
$AddWB_{BCH}$	BCH number of total addresses with balance	Log-difference	Crypto Market-Related
$\rho_1(VOC_{BTC},r_{BTC})$	BTC log-ret BTC volume on chain	1-day EMA linear correlation	Crypto Market-Related

Tag	Variable(s)	Transformation	Group
$\rho_1(\mathrm{VOC}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET volume on chain	1-day EMA linear correlation	Crypto Market-Related
$\rho_2(\mathrm{VOC}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC volume on chain	2-day EMA linear correlation	Crypto Market-Related
$\rho_2(\mathrm{VOC}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET volume on chain	2-day EMA linear correlation	Crypto Market-Related
$\rho_7(\mathrm{VOC}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC volume on chain	7-day EMA linear correlation	Crypto Market-Related
$\rho_7(\mathrm{VOC}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET volume on chain	7-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(\mathrm{VOC}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC volume on chain	14-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(\mathrm{VOC}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET volume on chain	14-day EMA linear correlation	Crypto Market-Related
$g_1(\mathrm{VOC}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC volume on chain	1-day EMA Gerber correlation	Crypto Market-Related
$g_1(\mathrm{VOC}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET volume on chain	1-day EMA Gerber correlation	Crypto Market-Related
$g_2(\mathrm{VOC}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC volume on chain	2-day EMA Gerber correlation	Crypto Market-Related
$g_2(\mathrm{VOC}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET volume on chain	2-day EMA Gerber correlation	Crypto Market-Related
$g_7(\mathrm{VOC}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC volume on chain	7-day EMA Gerber correlation	Crypto Market-Related
$g_7(\mathrm{VOC}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET volume on chain	7-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(\mathrm{VOC}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC volume on chain	14-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(\mathrm{VOC}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET volume on chain	14-day EMA Gerber correlation	Crypto Market-Related
$\rho_1(\mathrm{VOC}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC volume on chain	1-day EMA linear correlation	Crypto Market-Related
$\rho_1(\mathrm{VOC}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret ET volume on chain	1-day EMA linear correlation	Crypto Market-Related
$\rho_2(\mathrm{VOC}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC volume on chain	2-day EMA linear correlation	Crypto Market-Related
$\rho_2(\mathrm{VOC}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret ET volume on chain	2-day EMA linear correlation	Crypto Market-Related
$\rho_7(\mathrm{VOC}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC volume on chain	7-day EMA linear correlation	Crypto Market-Related
$\rho_7(\mathrm{VOC}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret ET volume on chain	7-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(\mathrm{VOC}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC volume on chain	14-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(\mathrm{VOC}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret ET volume on chain	14-day EMA linear correlation	Crypto Market-Related
$g_1(\mathrm{VOC}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC volume on chain	1-day EMA Gerber correlation	Crypto Market-Related

Tag	Variable(s)	Transformation	Group
$g_1(\mathrm{VOC}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret ET volume on chain	1-day EMA Gerber correlation	Crypto Market-Related
$g_2(\mathrm{VOC}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC volume on chain	2-day EMA Gerber correlation	Crypto Market-Related
$g_2(\mathrm{VOC}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret ET volume on chain	2-day EMA Gerber correlation	Crypto Market-Related
$g_7(\mathrm{VOC}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC volume on chain	7-day EMA Gerber correlation	Crypto Market-Related
$g_7(\mathrm{VOC}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret ET volume on chain	7-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(\mathrm{VOC}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC volume on chain	14-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(\mathrm{VOC}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret ET volume on chain	14-day EMA Gerber correlation	Crypto Market-Related
$\rho_1(\mathrm{AddWB}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC number of total addresses with balance with balance	1-day EMA linear correlation	Crypto Market-Related
$\rho_1(\mathrm{AddWB}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET number of total addresses with balance with balance	1-day EMA linear correlation	Crypto Market-Related
$\rho_1(\mathrm{AddWB}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC number of total addresses with balance with balance	1-day EMA linear correlation	Crypto Market-Related
$\rho_1(\mathrm{AddWB}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH number of total addresses with balance with balance	1-day EMA linear correlation	Crypto Market-Related
$\rho_2(\mathrm{AddWB}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC number of total addresses with balance with balance	2-day EMA linear correlation	Crypto Market-Related
$\rho_2(\mathrm{AddWB}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET number of total addresses with balance with balance	2-day EMA linear correlation	Crypto Market-Related
$\rho_2(\mathrm{AddWB}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC number of total addresses with balance with balance	2-day EMA linear correlation	Crypto Market-Related
$\rho_2(\mathrm{AddWB}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH number of total addresses with balance with balance	2-day EMA linear correlation	Crypto Market-Related
$\rho_7(\mathrm{AddWB}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC number of total addresses with balance with balance	7-day EMA linear correlation	Crypto Market-Related

Tag	Variable(s)	Transformation	Group
$\rho_7(\mathrm{AddWB}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET number of total addresses with balance with balance	7-day EMA linear correlation	Crypto Market-Related
$\rho_7(\mathrm{AddWB}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC number of total addresses with balance with balance	7-day EMA linear correlation	Crypto Market-Related
$\rho_7(\mathrm{AddWB}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH number of total addresses with balance with balance	7-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(\mathrm{AddWB}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC number of total addresses with balance with balance	14-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(\mathrm{AddWB}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET number of total addresses with balance with balance	14-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(\mathrm{AddWB}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC number of total addresses with balance with balance	14-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(\mathrm{AddWB}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH number of total addresses with balance with balance	14-day EMA linear correlation	Crypto Market-Related
$g_1(\mathrm{AddWB}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC number of total addresses with balance with balance	1-day EMA Gerber correlation	Crypto Market-Related
$g_1(\mathrm{AddWB}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET number of total addresses with balance with balance	1-day EMA Gerber correlation	Crypto Market-Related
$g_1(\mathrm{AddWB}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC number of total addresses with balance with balance	1-day EMA Gerber correlation	Crypto Market-Related
$g_1(\mathrm{AddWB}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH number of total addresses with balance with balance	1-day EMA Gerber correlation	Crypto Market-Related
$g_2(\mathrm{AddWB}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC number of total addresses with balance with balance	2-day EMA Gerber correlation	Crypto Market-Related
$g_2(\mathrm{AddWB}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET number of total addresses with balance with balance	2-day EMA Gerber correlation	Crypto Market-Related

Tag	Variable(s)	Transformation	Group
$g_2(\text{AddWB}_{\text{LTC}}, r_{\text{LTC}})$	LTC log-ret LTC number of total addresses with balance with balance	2-day EMA Gerber correlation	Crypto Market-Related
$g_2(\text{AddWB}_{\text{BCH}}, r_{\text{BCH}})$	BCH log-ret BCH number of total addresses with balance with balance	2-day EMA Gerber correlation	Crypto Market-Related
$g_7(\text{AddWB}_{\text{BTC}}, r_{\text{BTC}})$	BTC log-ret BTC number of total addresses with balance with balance	7-day EMA Gerber correlation	Crypto Market-Related
$g_7(\text{AddWB}_{\text{ETH}}, r_{\text{ETH}})$	ETH log-ret ET number of total addresses with balance with balance	7-day EMA Gerber correlation	Crypto Market-Related
$g_7(\text{AddWB}_{\text{LTC}}, r_{\text{LTC}})$	LTC log-ret LTC number of total addresses with balance with balance	7-day EMA Gerber correlation	Crypto Market-Related
$g_7(\text{AddWB}_{\text{BCH}}, r_{\text{BCH}})$	BCH log-ret BCH number of total addresses with balance with balance	7-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(\text{AddWB}_{\text{BTC}}, r_{\text{BTC}})$	BTC log-ret BTC number of total addresses with balance with balance	14-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(\text{AddWB}_{\text{ETH}}, r_{\text{ETH}})$	ETH log-ret ET number of total addresses with balance with balance	14-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(\text{AddWB}_{\text{LTC}}, r_{\text{LTC}})$	LTC log-ret LTC number of total addresses with balance with balance	14-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(\text{AddWB}_{\text{BCH}}, r_{\text{BCH}})$	BCH log-ret BCH number of total addresses with balance with balance	14-day EMA Gerber correlation	Crypto Market-Related
$\rho_1(\text{HR}_{\text{BTC}}, r_{\text{BTC}})$	BTC log-ret BTC hash rate	1-day EMA linear correlation	Crypto Market-Related
$\rho_1(\text{HR}_{\text{ETH}}, r_{\text{ETH}})$	ETH log-ret ET hash rate	1-day EMA linear correlation	Crypto Market-Related
$\rho_2(\text{HR}_{\text{BTC}}, r_{\text{BTC}})$	BTC log-ret BTC hash rate	2-day EMA linear correlation	Crypto Market-Related
$\rho_2(\text{HR}_{\text{ETH}}, r_{\text{ETH}})$	ETH log-ret ET hash rate	2-day EMA linear correlation	Crypto Market-Related
$\rho_7(\text{HR}_{\text{BTC}}, r_{\text{BTC}})$	BTC log-ret BTC hash rate	7-day EMA linear correlation	Crypto Market-Related

Tag	Variable(s)	Transformation	Group
$\rho_7(\mathrm{HR}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET hash rate	7-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(\mathrm{HR}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC hash rate	14-day EMA linear correlation	Crypto Market-Related
$\rho_{14}(\mathrm{HR}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET hash rate	14-day EMA linear correlation	Crypto Market-Related
$g_1(\mathrm{HR}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC hash rate	1-day EMA Gerber correlation	Crypto Market-Related
$g_1(\mathrm{HR}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET hash rate	1-day EMA Gerber correlation	Crypto Market-Related
$g_2(\mathrm{HR}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC hash rate	2-day EMA Gerber correlation	Crypto Market-Related
$g_2(\mathrm{HR}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET hash rate	2-day EMA Gerber correlation	Crypto Market-Related
$g_7(\mathrm{HR}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC hash rate	7-day EMA Gerber correlation	Crypto Market-Related
$g_7(\mathrm{HR}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET hash rate	7-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(\mathrm{HR}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC hash rate	14-day EMA Gerber correlation	Crypto Market-Related
$g_{14}(\mathrm{HR}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET hash rate	14-day EMA Gerber correlation	Crypto Market-Related
$\mathrm{RF}_1(\mathrm{BTC})$	BTC log-ret BTC Volatility	Time series regression forecast l = 1	Crypto Market-Related
$\mathrm{RF}_2(\mathrm{BTC})$	BTC log-ret BTC Volatility	Time series regression forecast l = 2	Crypto Market-Related
$\mathrm{RF}_7(\mathrm{BTC})$	BTC log-ret BTC Volatility	Time series regression forecast l = 7	Crypto Market-Related
$\mathrm{RF}_{14}(\mathrm{BTC})$	BTC log-ret BTC Volatility	Tme series regression forecast l = 14	Crypto Market-Related
$\mathrm{RF}_1(\mathrm{ETH})$	ETH log-ret ETH Volatility	Time series regression forecast l = 1	Crypto Market-Related
$\mathrm{RF}_2(\mathrm{ETH})$	ETH log-ret ETH Volatility	Time series regression forecast l = 2	Crypto Market-Related
$\mathrm{RF}_7(\mathrm{ETH})$	ETH log-ret ETH Volatility	Time series regression forecast l = 7	Crypto Market-Related
$\mathrm{RF}_{14}(\mathrm{ETH})$	ETH log-ret ETH Volatility	Time series regression forecast l = 14	Crypto Market-Related
$\mathrm{RF}_1(\mathrm{XRP})$	XRP log-ret XRP Volatility	Time series regression forecast l = 1	Crypto Market-Related
$\mathrm{RF}_2(\mathrm{XRP})$	XRP log-ret XRP Volatility	Time series regression forecast l = 2	Crypto Market-Related
$\mathrm{RF}_7(\mathrm{XRP})$	XRP log-ret XRP Volatility	Time series regression forecast l = 7	Crypto Market-Related
$\mathrm{RF}_{14}(\mathrm{XRP})$	XRP log-ret XRP Volatility	Time series regression forecast l =14	Crypto Market-Related
$\mathrm{RF}_1(\mathrm{LTC})$	LTC log-ret LTC Volatility	Time series regression forecast l = 1	Crypto Market-Related

Tag	Variable(s)	Transformation	Group
$RF_2(LTC)$	LTC log-ret LTC Volatility	Time series regression forecast l = 2	Crypto Market-Related
$RF_7(LTC)$	LTC log-ret LTC Volatility	Time series regression forecast l = 7	Crypto Market-Related
$RF_{14}(LTC)$	LTC log-ret LTC Volatility	Time series regression forecast l = 14	Crypto Market-Related
$RF_1(BCH)$	BCH log-ret BCH Volatility	Time series regression forecast l =1	Crypto Market-Related
$RF_2(BCH)$	BCH log-ret BCH Volatility	Time series regression forecast l = 2	Crypto Market-Related
$RF_7(BCH)$	BCH log-ret BCH Volatility	Time series regression forecast l = 7	Crypto Market-Related
$RF_{14}(BCH)$	BCH log-ret BCH Volatility	Time series regression forecast l = 14	Crypto Market-Related
RSI(BTC)	BTC Price	RSI	Crypto Market-Related
MACDS(BTC)	BTC Price	MACD minus signal	Crypto Market-Related
RSI(ETH)	ETH Price	RSI	Crypto Market-Related
MACDS(ETH)	ETH Price	MACD minus signal	Crypto Market-Related
RSI(XRP)	XRP Price	RSI	Crypto Market-Related
MACDS(XRP)	XRP Price	MACD minus signal	Crypto Market-Related
RSI(LTC)	LTC Price	RSI	Crypto Market-Related
MACDS(LTC)	LTC Price	MACD minus signal	Crypto Market-Related
RSI(BCH)	BCH Price	RSI	Crypto Market-Related
MACDS(BCH)	BCH Price	MACD minus signal	Crypto Market-Related
$\rho_1(r_{ETH}, r_{BTC})$	BTC log-ret ETH log-ret	1-day linear correlation	Crypto Market-Related
$\rho_1(r_{XRP}, r_{BTC})$	BTC log-ret XRP log-ret	1-day linear correlation	Crypto Market-Related
$\rho_1(r_{LTC}, r_{BTC})$	BTC log-ret LTC log-ret	1-day linear correlation	Crypto Market-Related
$\rho_1(r_{BCH}, r_{BTC})$	BTC log-ret BCH log-ret	1-day linear correlation	Crypto Market-Related
$\rho_2(r_{ETH}, r_{BTC})$	BTC log-ret ETH log-ret	2-day linear correlation	Crypto Market-Related
$\rho_2(r_{XRP}, r_{BTC})$	BTC log-ret XRP log-ret	2-day linear correlation	Crypto Market-Related
$\rho_2(r_{LTC}, r_{BTC})$	BTC log-ret LTC log-ret	2-day linear correlation	Crypto Market-Related
$\rho_2(r_{BCH}, r_{BTC})$	BTC log-ret BCH log-ret	2-day linear correlation	Crypto Market-Related
$\rho_7(r_{ETH}, r_{BTC})$	BTC log-ret ETH log-ret	7-day linear correlation	Crypto Market-Related
$\rho_7(r_{XRP}, r_{BTC})$	BTC log-ret XRP log-ret	7-day linear correlation	Crypto Market-Related
$\rho_7(r_{LTC}, r_{BTC})$	BTC log-ret LTC log-ret	7-day linear correlation	Crypto Market-Related

Tag	Variable(s)	Transformation	Group
$\rho_7(r_{\text{BCH}}, r_{\text{BTC}})$	BTC log-ret BCH log-ret	7-day linear correlation	Crypto Market-Related
$\rho_{14}(r_{\text{ETH}}, r_{\text{BTC}})$	BTC log-ret ETH log-ret	14-day linear correlation	Crypto Market-Related
$\rho_{14}(r_{\text{XRP}}, r_{\text{BTC}})$	BTC log-ret XRP log-ret	14-day linear correlation	Crypto Market-Related
$\rho_{14}(r_{\text{LTC}}, r_{\text{BTC}})$	BTC log-ret LTC log-ret	14-day linear correlation	Crypto Market-Related
$\rho_{14}(r_{\text{BCH}}, r_{\text{BTC}})$	BTC log-ret BCH log-ret	14-day linear correlation	Crypto Market-Related
$g_1(r_{\text{ETH}}, r_{\text{BTC}})$	BTC log-ret ETH log-ret	1-day linear correlation	Crypto Market-Related
$g_1(r_{\text{XRP}}, r_{\text{BTC}})$	BTC log-ret XRP log-ret	1-day linear correlation	Crypto Market-Related
$g_1(r_{\text{LTC}}, r_{\text{BTC}})$	BTC log-ret LTC log-ret	1-day linear correlation	Crypto Market-Related
$g_1(r_{\text{BCH}}, r_{\text{BTC}})$	BTC log-ret BCH log-ret	1-day linear correlation	Crypto Market-Related
$g_2(r_{\text{ETH}}, r_{\text{BTC}})$	BTC log-ret ETH log-ret	2-day linear correlation	Crypto Market-Related
$g_2(r_{\text{XRP}}, r_{\text{BTC}})$	BTC log-ret XRP log-ret	2-day linear correlation	Crypto Market-Related
$g_2(r_{\text{LTC}}, r_{\text{BTC}})$	BTC log-ret LTC log-ret	2-day linear correlation	Crypto Market-Related
$g_2(r_{\text{BCH}}, r_{\text{BTC}})$	BTC log-ret BCH log-ret	2-day linear correlation	Crypto Market-Related
$g_7(r_{\text{ETH}}, r_{\text{BTC}})$	BTC log-ret ETH log-ret	7-day linear correlation	Crypto Market-Related
$g_7(r_{\text{XRP}}, r_{\text{BTC}})$	BTC log-ret XRP log-ret	7-day linear correlation	Crypto Market-Related
$g_7(r_{\text{LTC}}, r_{\text{BTC}})$	BTC log-ret LTC log-ret	7-day linear correlation	Crypto Market-Related
$g_7(r_{\text{BCH}}, r_{\text{BTC}})$	BTC log-ret BCH log-ret	7-day linear correlation	Crypto Market-Related
$g_{14}(r_{\text{ETH}}, r_{\text{BTC}})$	BTC log-ret ETH log-ret	14-day linear correlation	Crypto Market-Related
$g_{14}(r_{\text{XRP}}, r_{\text{BTC}})$	BTC log-ret XRP log-ret	14-day linear correlation	Crypto Market-Related
$g_{14}(r_{\text{LTC}}, r_{\text{BTC}})$	BTC log-ret LTC log-ret	14-day linear correlation	Crypto Market-Related
$g_{14}(r_{\text{BCH}}, r_{\text{BTC}})$	BTC log-ret BCH log-ret	14-day linear correlation	Crypto Market-Related
TM(BTC)	BTC on chain volume with balance BTC prices	First difference of ratio	Crypto Market-Related
TM(ETH)	ETH on chain volume with balance ETH prices	First difference of ratio	Crypto Market-Related

Tag	Variable(s)	Transformation	Group
TM(LTC)	LTC on chain volume with balance LTC prices	First difference of ratio	Crypto Market-Related
TM(BCH)	BCH number of total addresses with balance BCH prices	First difference of ratio	Crypto Market-Related
AM(BTC)	BTC number of total addresses BTC prices	Log-difference of ratio	Crypto Market-Related
AM(ETH)	ETH number of total addresses ETH prices	Log-difference of ratio	Crypto Market-Related
AM(LTC)	LTC number of total addresses LTC prices	Log-difference of ratio	Crypto Market-Related
AM(BCH)	BCH number of total addresses BCH prices	Log-difference of ratio	Crypto Market-Related
UM(BTC)	BTC number of total addresses with balance BTC prices	Log-difference of ratio	Crypto Market-Related
UM(ETH)	ETH number of total addresses with balance ETH prices	Log-difference of ratio	Crypto Market-Related
UM(LTC)	LTC number of total addresses with balance LTC prices	Log-difference of ratio	Crypto Market-Related
UM(BCH)	BCH number of total addresses with balance BCH prices	Log-difference of ratio	Crypto Market-Related
$AMIHUD_{BTC}$	BTC absolute log-ret BTC volume	Ratio	Crypto Market-Related
$AMIHUD_{ETH}$	ETH absolute log-ret ETH volume	Ratio	Crypto Market-Related
$AMIHUD_{XRP}$	XRP absolute log-ret XRP volume	Ratio	Crypto Market-Related
$AMIHUD_{LTC}$	LTC absolute log-ret LTC volume	Ratio	Crypto Market-Related
$AMIHUD_{BCH}$	BCH absolute log-ret BCH volume	Ratio	Crypto Market-Related
GT_{BTC}	BTC Google Index	First difference	Sentiment
GT_{ETH}	ETH Google Index	First difference	Sentiment
GT_{XRP}	XRP Google Index	First difference	Sentiment
GT_{LTC}	LTC Google Index	First difference	Sentiment
GT_{BCH}	BCH Google Index	First difference	Sentiment
$\rho_1(GT_{BTC}, r_{BTC})$	BTC log-ret BTC Google	1-day EMA linear correlation	Sentiment
$\rho_1(GT_{ETH}, r_{ETH})$	ETH log-ret ET Google	1-day EMA linear correlation	Sentiment

Tag	Variable(s)	Transformation	Group
$\rho_1(\mathrm{GT}_{\mathrm{XRP}}, r_{\mathrm{XRP}})$	XRP log-ret XRP Google	1-day EMA linear correlation	Sentiment
$\rho_1(\mathrm{GT}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC Google	1-day EMA linear correlation	Sentiment
$\rho_1(\mathrm{GT}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH Google	1-day EMA linear correlation	Sentiment
$\rho_2(\mathrm{GT}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC Google	2-day EMA linear correlation	Sentiment
$\rho_2(\mathrm{GT}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET Google	2-day EMA linear correlation	Sentiment
$\rho_2(\mathrm{GT}_{\mathrm{XRP}}, r_{\mathrm{XRP}})$	XRP log-ret XRP Google	2-day EMA linear correlation	Sentiment
$\rho_2(\mathrm{GT}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC Google	2-day EMA linear correlation	Sentiment
$\rho_2(\mathrm{GT}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH Google	2-day EMA linear correlation	Sentiment
$\rho_7(\mathrm{GT}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC Google	7-day EMA linear correlation	Sentiment
$\rho_7(\mathrm{GT}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET Google	7-day EMA linear correlation	Sentiment
$\rho_7(\mathrm{GT}_{\mathrm{XRP}}, r_{\mathrm{XRP}})$	XRP log-ret XRP Google	7-day EMA linear correlation	Sentiment
$\rho_7(\mathrm{GT}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC Google	7-day EMA linear correlation	Sentiment
$\rho_7(\mathrm{GT}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH Google	7-day EMA linear correlation	Sentiment
$\rho_{14}(\mathrm{GT}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC Google	14-day EMA linear correlation	Sentiment
$\rho_{14}(\mathrm{GT}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET Google	14-day EMA linear correlation	Sentiment
$\rho_{14}(\mathrm{GT}_{\mathrm{XRP}}, r_{\mathrm{XRP}})$	XRP log-ret XRP Google	14-day EMA linear correlation	Sentiment
$\rho_{14}(\mathrm{GT}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC Google	14-day EMA linear correlation	Sentiment
$\rho_{14}(\mathrm{GT}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH Google	14-day EMA linear correlation	Sentiment
$g_1(\mathrm{GT}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC Google	1-day EMA Gerber correlation	Sentiment
$g_1(\mathrm{GT}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET Google	1-day EMA Gerber correlation	Sentiment
$g_1(\mathrm{GT}_{\mathrm{XRP}}, r_{\mathrm{XRP}})$	XRP log-ret XRP Google	1-day EMA Gerber correlation	Sentiment
$g_1(\mathrm{GT}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC Google	1-day EMA Gerber correlation	Sentiment
$g_1(\mathrm{GT}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH Google	1-day EMA Gerber correlation	Sentiment

Tag	Variable(s)	Transformation	Group
$g_2(\mathrm{GT}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC Google	2-day EMA Gerber correlation	Sentiment
$g_2(\mathrm{GT}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET Google	2-day EMA Gerber correlation	Sentiment
$g_2(\mathrm{GT}_{\mathrm{XRP}}, r_{\mathrm{XRP}})$	XRP log-ret XRP Google	2-day EMA Gerber correlation	Sentiment
$g_2(\mathrm{GT}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC Google	2-day EMA Gerber correlation	Sentiment
$g_2(\mathrm{GT}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH Google	2-day EMA Gerber correlation	Sentiment
$g_7(\mathrm{GT}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC Google	7-day EMA Gerber correlation	Sentiment
$g_7(\mathrm{GT}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET Google	7-day EMA Gerber correlation	Sentiment
$g_7(\mathrm{GT}_{\mathrm{XRP}}, r_{\mathrm{XRP}})$	XRP log-ret XRP Google	7-day EMA Gerber correlation	Sentiment
$g_7(\mathrm{GT}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC Google	7-day EMA Gerber correlation	Sentiment
$g_7(\mathrm{GT}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH Google	7-day EMA Gerber correlation	Sentiment
$g_{14}(\mathrm{GT}_{\mathrm{BTC}}, r_{\mathrm{BTC}})$	BTC log-ret BTC Google	14-day EMA Gerber correlation	Sentiment
$g_{14}(\mathrm{GT}_{\mathrm{ETH}}, r_{\mathrm{ETH}})$	ETH log-ret ET Google	14-day EMA Gerber correlation	Sentiment
$g_{14}(\mathrm{GT}_{\mathrm{XRP}}, r_{\mathrm{XRP}})$	XRP log-ret XRP Google	14-day EMA Gerber correlation	Sentiment
$g_{14}(\mathrm{GT}_{\mathrm{LTC}}, r_{\mathrm{LTC}})$	LTC log-ret LTC Google	14-day EMA Gerber correlation	Sentiment
$g_{14}(\mathrm{GT}_{\mathrm{BCH}}, r_{\mathrm{BCH}})$	BCH log-ret BCH Google	14-day EMA Gerber correlation	Sentiment
GOLD	Gold log-ret	Log-difference	Financial Market
EURUSD	EURUSD log-ret	Log-difference	Financial Market
JPYUSD	JPYUSD log-ret	Log-difference	Financial Market
CNYUSD	CNYUSD log-ret	Log-difference	Financial Market
VIX	VIX log-ret	Log-difference	Financial Market
SP500	SP500 log-ret	Log-difference	Financial Market
NASDAQ	NASDAQ log-ret	Log-difference	Financial Market
T10Y3M	10 years minus 3 months US treasury yields	First difference	Financial Market
WTI	WTI	First difference	Financial Market
$\rho_1(\mathrm{GOLD}, r_{\mathrm{BTC}})$	BTC log-ret gold log-ret	1-day EMA Gerber correlation	Financial Market
$\rho_2(\mathrm{GOLD}, r_{\mathrm{BTC}})$	BTC log-ret gold log-ret	2-day EMA Gerber correlation	Financial Market
$\rho_7(\mathrm{GOLD}, r_{\mathrm{BTC}})$	BTC log-ret gold log-ret	7-day EMA Gerber correlation	Financial Market

Tag	Variable(s)	Transformation	Group
ρ_{14}(GOLD,r_{BTC})	BTC log-ret gold log-ret	14-day EMA Gerber correlation	Financial Market
g_1(GOLD,r_{BTC})	BTC log-ret gold log-ret	1-day EMA Gerber correlation	Financial Market
g_2(GOLD,r_{BTC})	BTC log-ret gold log-ret	2-day EMA Gerber correlation	Financial Market
g_7(GOLD,r_{BTC})	BTC log-ret gold log-ret	7-day EMA Gerber correlation	Financial Market
g_{14}(GOLD,r_{BTC})	BTC log-ret gold log-ret	14-day EMA Gerber correlation	Financial Market
ρ_1(EURUSD,r_{BTC})	BTC log-ret EURUSD log-ret	1-day EMA Gerber correlation	Financial Market
ρ_2(EURUSD,r_{BTC})	BTC log-ret EURUSD log-ret	2-day EMA Gerber correlation	Financial Market
ρ_7(EURUSD,r_{BTC})	BTC log-ret EURUSD log-ret	7-day EMA Gerber correlation	Financial Market
ρ_{14}(EURUSD,r_{BTC})	BTC log-ret EURUSD log-ret	14-day EMA Gerber correlation	Financial Market
g_1(EURUSD,r_{BTC})	BTC log-ret EURUSD log-ret	1-day EMA Gerber correlation	Financial Market
g_2(EURUSD,r_{BTC})	BTC log-ret EURUSD log-ret	2-day EMA Gerber correlation	Financial Market
g_7(EURUSD,r_{BTC})	BTC log-ret EURUSD log-ret	7-day EMA Gerber correlation	Financial Market
g_{14}(EURUSD,r_{BTC})	BTC log-ret EURUSD log-ret	14-day EMA Gerber correlation	Financial Market
ρ_1(JPYUSD,r_{BTC})	BTC log-ret JPYUSD log-ret	1-day EMA Gerber correlation	Financial Market
ρ_2(JPYUSD,r_{BTC})	BTC log-ret JPYUSD log-ret	2-day EMA Gerber correlation	Financial Market
ρ_7(JPYUSD,r_{BTC})	BTC log-ret JPYUSD log-ret	7-day EMA Gerber correlation	Financial Market
ρ_{14}(JPYUSD,r_{BTC})	BTC log-ret JPYUSD log-ret	14-day EMA Gerber correlation	Financial Market
g_1(JPYUSD,r_{BTC})	BTC log-ret JPYUSD log-ret	1-day EMA Gerber correlation	Financial Market
g_2(JPYUSD,r_{BTC})	BTC log-ret JPYUSD log-ret	2-day EMA Gerber correlation	Financial Market
g_7(JPYUSD,r_{BTC})	BTC log-ret JPYUSD log-ret	7-day EMA Gerber correlation	Financial Market
g_{14}(JPYUSD,r_{BTC})	BTC log-ret JPYUSD log-ret	14-day EMA Gerber correlation	Financial Market
ρ_1(CNYUSD,r_{BTC})	BTC log-ret CNYUSD log-ret	1-day EMA Gerber correlation	Financial Market
ρ_2(CNYUSD,r_{BTC})	BTC log-ret CNYUSD log-ret	2-day EMA Gerber correlation	Financial Market
ρ_7(CNYUSD,r_{BTC})	BTC log-ret CNYUSD log-ret	7-day EMA Gerber correlation	Financial Market

Tag	Variable(s)	Transformation	Group
ρ_{14}(CNYUSD,r_{BTC})	BTC log-ret CNYUSD log-ret	14-day EMA Gerber correlation	Financial Market
g_1(CNYUSD,r_{BTC})	BTC log-ret CNYUSD log-ret	1-day EMA Gerber correlation	Financial Market
g_2(CNYUSD,r_{BTC})	BTC log-ret CNYUSD log-ret	2-day EMA Gerber correlation	Financial Market
g_7(CNYUSD,r_{BTC})	BTC log-ret CNYUSD log-ret	7-day EMA Gerber correlation	Financial Market
g_{14}(CNYUSD,r_{BTC})	BTC log-ret CNYUSD log-ret	14-day EMA Gerber correlation	Financial Market
ρ_1(VIX,r_{BTC})	BTC log-ret VIX log-ret	1-day EMA Gerber correlation	Financial Market
ρ_2(VIX,r_{BTC})	BTC log-ret VIX log-ret	2-day EMA Gerber correlation	Financial Market
ρ_7(VIX,r_{BTC})	BTC log-ret VIX log-ret	7-day EMA Gerber correlation	Financial Market
ρ_{14}(VIX,r_{BTC})	BTC log-ret VIX log-ret	14-day EMA Gerber correlation	Financial Market
g_1(VIX,r_{BTC})	BTC log-ret VIX log-ret	1-day EMA Gerber correlation	Financial Market
g_2(VIX,r_{BTC})	BTC log-ret VIX log-ret	2-day EMA Gerber correlation	Financial Market
g_7(VIX,r_{BTC})	BTC log-ret VIX log-ret	7-day EMA Gerber correlation	Financial Market
g_{14}(VIX,r_{BTC})	BTC log-ret VIX log-ret	14-day EMA Gerber correlation	Financial Market
ρ_1(SP500,r_{BTC})	BTC log-ret SP500 log-ret	1-day EMA Gerber correlation	Financial Market
ρ_2(SP500,r_{BTC})	BTC log-ret SP500 log-ret	2-day EMA Gerber correlation	Financial Market
ρ_7(SP500,r_{BTC})	BTC log-ret SP500 log-ret	7-day EMA Gerber correlation	Financial Market
ρ_{14}(SP500,r_{BTC})	BTC log-ret SP500 log-ret	14-day EMA Gerber correlation	Financial Market
g_1(SP500,r_{BTC})	BTC log-ret SP500 log-ret	1-day EMA Gerber correlation	Financial Market
g_2(SP500,r_{BTC})	BTC log-ret SP500 log-ret	2-day EMA Gerber correlation	Financial Market
g_7(SP500,r_{BTC})	BTC log-ret SP500 log-ret	7-day EMA Gerber correlation	Financial Market
g_{14}(SP500,r_{BTC})	BTC log-ret SP500 log-ret	14-day EMA Gerber correlation	Financial Market
ρ_1(NASDAQ,r_{BTC})	BTC log-ret NASDAQ log-ret	1-day EMA Gerber correlation	Financial Market
ρ_2(NASDAQ,r_{BTC})	BTC log-ret NASDAQ log-ret	2-day EMA Gerber correlation	Financial Market
ρ_7(NASDAQ,r_{BTC})	BTC log-ret NASDAQ log-ret	7-day EMA Gerber correlation	Financial Market

Tag	Variable(s)	Transformation	Group
ρ_{14}(NASDAQ,r_{BTC})	BTC log-ret NASDAQ log-ret	14-day EMA Gerber correlation	Financial Market
g_1(NASDAQ,r_{BTC})	BTC log-ret NASDAQ log-ret	1-day EMA Gerber correlation	Financial Market
g_2(NASDAQ,r_{BTC})	BTC log-ret NASDAQ log-ret	2-day EMA Gerber correlation	Financial Market
g_7(NASDAQ,r_{BTC})	BTC log-ret NASDAQ log-ret	7-day EMA Gerber correlation	Financial Market
g_{14}(NASDAQ,r_{BTC})	BTC log-ret NASDAQ log-ret	14-day EMA Gerber correlation	Financial Market
ρ_1 (T10Y3M, r_{BTC})	BTC log-ret T10Y3M	1-day EMA Gerber correlation	Financial Market
ρ_2 (T10Y3M, r_{BTC})	BTC log-ret T10Y3M	2-day EMA Gerber correlation	Financial Market
ρ_7 (T10Y3M, r_{BTC})	BTC log-ret T10Y3M	7-day EMA Gerber correlation	Financial Market
ρ_{14} (T10Y3M, r_{BTC})	BTC log-ret T10Y3M	14-day EMA Gerber correlation	Financial Market
g_1 (T10Y3M, r_{BTC})	BTC log-ret T10Y3M	1-day EMA Gerber correlation	Financial Market
g_2 (T10Y3M, r_{BTC})	BTC log-ret T10Y3M	2-day EMA Gerber correlation	Financial Market
g_7 (T10Y3M, r_{BTC})	BTC log-ret T10Y3M	7-day EMA Gerber correlation	Financial Market
g_{14} (T10Y3M, r_{BTC})	BTC log-ret T10Y3M	14-day EMA Gerber correlation	Financial Market
ρ_1(WTI,r_{BTC})	BTC log-ret WTI	1-day EMA Gerber correlation	Financial Market
ρ_2(WTI,r_{BTC})	BTC log-ret WTI	2-day EMA Gerber correlation	Financial Market
ρ_7(WTI,r_{BTC})	BTC log-ret WTI	7-day EMA Gerber correlation	Financial Market
ρ_{14}(WTI,r_{BTC})	BTC log-ret WTI	14-day EMA Gerber correlation	Financial Market
g_1(WTI,r_{BTC})	BTC log-ret WTI	1-day EMA Gerber correlation	Financial Market
g_2(WTI,r_{BTC})	BTC log-ret WTI	2-day EMA Gerber correlation	Financial Market
g_7(WTI,r_{BTC})	BTC log-ret WTI	7-day EMA Gerber correlation	Financial Market
g_{14}(WTI,r_{BTC})	BTC log-ret WTI	14-day EMA Gerber correlation	Financial Market
EMA_1(VIX)	VIX log-ret	1-day EMA	Financial Market
EMA_2(VIX)	VIX log-ret	2-day EMA	Financial Market
EMA_7(VIX)	VIX log-ret	7-day EMA	Financial Market
EMA_{14}(VIX)	VIX log-ret	14-day EMA	Financial Market

Acknowledgements We acknowledge the University of Milano-Bicocca Crypto Asset Lab and Data Science Lab for supporting this work by providing data and computational resources.

Funding Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement. Petter N. Kolm and Erik Lindström were partly supported by the Knut and Alice Wallenberg Foundation under grant KAW 2020.0280.

Declarations

Conflict of interest We have no conflicts of interest to disclose.

References

Aalborg, H. A., Molnár, P., & de Vries, J. E. (2019). What can explain the price, volatility and trading volume of Bitcoin? *Finance Research Letters, 29*, 255–265.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723.

Alexander, C., & Dakos, M. (2020). A critical investigation of cryptocurrency data and analysis. *Quantitative Finance, 20*(2), 173–188.

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets, 5*(1), 31–56.

Appel, G. (2005). *Technical analysis: Power tools for active investors*. FT Press.

Ardia, D., & Bluteau, K., Rüede M. (2019). Regime changes in Bitcoin GARCH volatility dynamics. *Finance Research Letters, 29*, 266–271.

Auer, R. , Cornelli, G. , Doerr, S. , Frost, J. , Gambacorta, L. (2022). Crypto trading and bitcoin prices: E vidence from a new database of retail adoption (Tech. Rep.). Bank for International Settlements .

Babu, A., Levine, A., Ooi, Y. H., Pedersen, L. H., & Stamelos, E. (2020). *Trends everywhere. Journal of Investment Management, 18*(1), 52–68.

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies, 21*(2), 785–818.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance, 61*(4), 1645–1680.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, 49*(3), 307–343.

Barucci, E., Moncayo, G. G., & Marazzina, D. (2022). Cryptocurrencies and stablecoins: A high-frequency analysis. *Digital Finance, 4*(2), 217–239.

Baur, D. G., Hong, K., & Lee, A. D. (2018). Bitcoin: M edium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money, 54*, 177–189.

Bemporad, A., Breschi, V., Piga, D., & Boyd, S. P. (2018). Fitting jump models. *Automatica, 96*, 11–21.

Bianchi, D., Dickerson, A. (2019). Trading volume in cryptocurrency markets . Available at SSRN 3239670.

Bianchi, D. (2020). Cryptocurrencies as an asset class? A n empirical assessment. *The Journal of Alternative Investments, 23*(2), 162–179.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy, 100*(5), 992–1026.

Bulla, J. (2011). Hidden Markov models with t components. Increased persistence and other aspects. *Quantitative Finance, 11*(3), 459–475.

Cappiello, L., Engle, R. F., & Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics, 4*(4), 537–572.

Catania, L., Grassi, S., & Ravazzolo, F. (2019). Forecasting cryptocurrencies under model and parameter instability. *International Journal of Forecasting, 35*, 485–501.

Chaim, P., & Laurini, M. P. (2018). Volatility and return jumps in Bitcoin. *Economics Letters, 173*, 158–163.

Cheah, J.E.-T., Luo, D., Zhang, Z., & Sung, M.-C. (2020). Predictability of Bitcoin returns. *The European Journal of Finance, 28*(1), 66–85.

Cong, L.W. , Karolyi, G.A. , Tang, K. , Zhao, W. (2021).Value premium, network adoption, and factor pricing of crypto assets. Working Paper.

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance, 1*(2), 223.
Cortese, F. P., Kolm, P. N., & Lindström, E. (2023). Generalized information criteria for sparse statistical jump models. In P. Linde (Ed.), Symposium i anvendt statistik, Vol 44. Copenhagen: Copenhagen Business School. http://www.statistiksymposium.dk/Symposium%20i%20anvendt%20stalistik%202023_Web.pdf
Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market underand overreactions. *The Journal of Finance, 53*(6), 1839–1885.
De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance, 45*(2), 379–395.
De Bandt, O., Hartmann, P. (2000). Systemic risk: A survey. European Central Bank Working Paper35.
Erb, C. B., Harvey, C. R., & Viskanta, T. E. (1994). Forecasting international equity correlations. *Financial Analysts Journal, 50*(6), 32–45.
Fan, Y., & Tang, C. Y. (2013). Tuning parameter selection in high dimensional penalized likelihood. *Journal of the Royal Statistical Society : Series B (Statistical Methodology), 75*(3), 531–552.
Figà-Talamanca, G., Focardi, S., & Patacca, M. (2021). Regime switches and commonalities of the cryptocurrencies asset class. *The North American Journal of Economics and Finance, 57*, 101425.
Figa-Talamanca, G., & Patacca, M. (2019). Does market attention affect Bitcoin returns and volatility? *Decisions in Economics and Finance, 42*(1), 135–155.
Gerber, S., Markowitz, H. M., Ernst, P. A., Miao, Y., Javid, B., & Sargen, P. (2022). The Gerber statistic: A robust co-movement measure for portfolio optimization. *The Journal of Portfolio Management, 48*(3), 87–102.
Hackethal, A., Hanspal, T., Lammer, D. M., & Rink, K. (2022). The characteristics and portfolio behavior of bitcoin investors: evidence from indirect cryptocurrency investments. *Review of Finance, 26*(4), 855–898.
Hamilton, J. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica, 57*, 357–384.
Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of finance, 54*(6), 2143–2184.
Karniol-Tambour, K. , Tan, R. , Tsarapkina, D. , Sondheimer, J. , Barnes, W. (2022). The evolution of institutional investors' exposure to cryptocurrencies and blockchain technologies (Tech. Rep.). Bridgewater Associates, LP.
Kogan, S. , Makarov, I. , Niessner, M. , Schoar, A. (2022). Are cryptos different? Evidence from retail trading. Available at SSRN 4289513.
Koki, C., Leonardos, S., & Piliouras, G. (2022). Exploring the predictability of cryptocurrencies via Bayesian hidden Markov models. *Research in International Business and Finance, 59*, 101554.
Kristoufek, L. (2013). Bitcoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the internet era. *Scientific Reports, 3*(1), 3415.
Kristoufek, L. (2015). What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. *PloS One, 10*, e0123923.
Lindström, E., Madsen, H., & Nielsen, J. N. (2015). *Statistics for finance: Texts in statistical science*. Chapman and Hall/CRC.
Liu, Y., & Tsyvinski, A. (2021). Risks and returns of cryptocurrency. *The Review of Financial Studies, 34*(6), 2689–2727.
Liu, Y., Tsyvinski, A., & Wu, X. (2022). Common risk factors in cryptocurrency. *The Journal of Finance, 77*(2), 1133–1177.
Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics, 104*(2), 228–250.
Nystrup, P., Kolm, P. N., & Lindström, E. (2020). Greedy online classification of persistent market states using realized intraday volatility features. *The Journal of Financial Data Science, 2*(3), 25–39.
Nystrup, P., Kolm, P. N., & Lindström, E. (2021). Feature selection in jump models. *Expert Systems with Applications, 184*, 115558.
Nystrup, P., Lindström, E., & Madsen, H. (2020). Learning hidden Markov models with persistent states by penalizing jumps. *Expert Systems with Applications, 150*, 113307.
Pele, D.T. , Wesselhöfft, N. , Härdle, W.K. , Kolossiatis, M. , Yatracos, Y.G. (2021). Are cryptos becoming alternative assets? The European Journal of Finance 1–42.
Pennoni, F., Bartolucci, F., Forte, G., & Ametrano, F. (2021). Exploring the dependencies among main cryptocurrency log-returns: A hidden Markov model. *Economic Notes, 51*, e12193.

Rydén, T. (2008). EM versus Markov chain Monte Carlo for estimation of hidden Markov models: A computational perspective. *Bayesian Analysis, 3*(4), 659–688.

Rydén, T., Teräsvirta, T., & Åsbrink, S. (1998). Stylized facts of daily return series and the hidden Markov model. *Journal of Applied Econometrics, 13*(3), 217–244.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464.

Selmi, R., Mensi, W., Hammoudeh, S., & Bouoiyour, J. (2018). Is Bitcoin a hedge, a safe haven or a diversifier for oil price movements? A comparison with gold. *Energy Economics, 74*, 787–801.

Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance, 40*(3), 777–790.

Shen, D., Urquhart, A., & Wang, P. (2020). Forecasting the volatility of Bitcoin: The importance of jumps and structural breaks. *European Financial Management, 26*(5), 1294–1323.

Smales, L. A. (2022). Investor attention in cryptocurrency markets. *International Review of Financial Analysis, 79*, 101972.

Urquhart, A. (2018). What causes the attention of Bitcoin? *Economics Letters, 166*, 40–44.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*, 1–67.

Vigna, P. (2022). *Wall street takes lead in crypto investments*. The Wall Street Journal.

Wilder, J. W. (1978). *New concepts in technical trading systems*. Trend Research.

Witten, D. M., & Tibshirani, R. (2010). A framework for feature selection in clustering. *Journal of the American Statistical Association, 105*(490), 713–726.

Xiong, J., Liu, Q., & Zhao, L. (2020). A new method to verify Bitcoin bubbles based on the production cost. *North American Journal of Economics and Finance, 51*, 101095.

Yae, J., & Tian, G. Z. (2022). Out-of-sample forecasting of cryptocurrency returns: A comprehensive comparison of predictors and algorithms. *Physica A: Statistical Mechanics and its Applications, 598*, 127379.

Yonekura, S., Beskos, A., & Singh, S. S. (2021). Asymptotic analysis of model selection criteria for general hidden Markov models. *Stochastic Processes and their Applications, 132*, 164–191.

Zucchini, W., MacDonald, I. L., & Langrock, R. (2017). *Hidden Markov Models for Time Series: An Introduction Using R*. Boca Raton: FLCRC Press.