Soraia Santos
KPMG, Lisboa
soraiapsantos6@gmail.com
orcid.org/0009-0002-1944-1028
Helder Sebastião
Univ Coimbra, CeBER, Faculty of Economics
helderse@fe.uc.pt
orcid.org/0000-0002-5687-3818
Nuno Silva
Univ Coimbra, CeBER, Faculty of Economics
nunos@fe.uc.pt
orcid.org/0000-0002-5687-3818
DOI: https://doi.org/10.14195/2183-203X_59_1

Bitcoin and Main Altcoins: Causality and Trading Strategies

Bitcoin e Principais Altcoins: Causalidade e Estratégias de Negociação

Soraia Santos
Helder Sebastião
Nuno Silva

ABSTRACT
Using daily data from November 9, 2017 to December 31, 2022, this paper uses Granger causality in the mean and the distribution to investigate the transmission of information between return, volume, volatility, and illiquidity for Bitcoin and the nine most important altcoins in terms of market capitalization. Additionally, the forecastability of Bitcoin returns is examined using linear models with different predictor spaces estimated using LASSO and the performance of several trading strategies devised upon those forecasts is assessed. The causal relationships between returns, volumes and volatilities of Bitcoin and each altcoin are more evident in the left tail of the distribution, where Bitcoin acts mostly as a transmitter of information, and in the right tail for causality regarding illiquidity. In bullish markets, Bitcoin acts mostly as a receiver of information. The best Bitcoin trading strategy is based on the model which incorporates the information on all cryptocurrencies, exhibiting a cumulative return of 331% and an annualized Sharpe ratio of 94.59%, considering an enter/exit threshold of 0.25% and after 0.5% round-trip transaction costs. These results are statistically significant when compared with the buy-and-hold strategy, which renders a cumulative return of 121% and a Sharpe ratio of 64.74%. These results point out the importance of considering information from other cryptocurrencies to forecast and trade on Bitcoin.
Keywords: Cryptocurrencies, Granger causality, LASSO, trading strategies.

JEL Classification: G11; G15; G17

Acknowledgments: This work has been funded by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., Project UIDB/05037/2020 and Project RiskBigData PTDC/MAT-APL/1286/2021.

1. Introduction

Since Bitcoin's inception in 2008, cryptocurrencies have achieved an important role as an alternative means of payment to traditional currencies, and, most notably, as a means for highly speculative investments.

According to the CMVM (Comissão do Mercado de Valores Mobiliários, i.e., Portuguese Securities Market Commission, 2022), crypto-assets are "*digital representations of assets based on blockchain technology, not issued by a central bank, credit institution or electronic money institution and that can be used as a form of payment in a community that accepts it or have other purposes such as the attribution of the right to use of certain goods and services or to a financial return*".

Cryptocurrencies are the subject of hot debates. On the one hand, they are perceived by many as a key point of an ongoing digital revolution, where transparency and decentralization are highlighted. On the other hand, many others point out the risks associated with its speculative nature and the independence of accredited and reliable institutions to guarantee transactions. Nevertheless, it is remarkable that, several years after the launch of Bitcoin, the cryptocurrency market continues to grow, and Bitcoin prevails as a leader in terms of acceptance and market capitalization (Sebastião et al., 2021)

Recent studies have addressed various topics inherent to cryptocurrencies with the aim of better understanding this market. With the emergence of more and more altcoins (alternative cryptocurrencies to Bitcoin) thriving, a relevant topic that still raises questions in the literature is the causal relationship between these cryptocurrencies and Bitcoin, the oldest cryptocurrency and the one with the largest market capitalisation. As such, the first objective of this study is to contribute to this theme, by analysing the transmission of information between Bitcoin and nine major competing cryptocurrencies (Ethereum, Binance Coin, Ripple, Cardano, Dogecoin, Tron, Ethereum Classic, Litecoin and Chainlink), regarding their return, transaction volume, volatility and illiquidity. This study is conducted using Granger causality tests not only in the mean but also in all distribution support.

Secondly, the goal is to define various trading strategies for Bitcoin by forecasting its profitability and then evaluating its performance. To forecast Bitcoin returns, we consider not only past information about Bitcoin but also lagged information about other cryptocurrencies. Thus, the second objective of this study is to analyse various trading strategies, as well as understand whether the predictive power improves when other cryptocurrencies are added to the model and its impact on the trading strategies' performance.

The originality of this study comes from its overall framework. Although several studies tackle some issues dealt with here, we provide a coherent framework that considers several variables of different cryptocurrencies, considers not only causality in the mean but also in the distribution, uses LASSO to select dynamically the information set, makes forecasts pooling the models and assess statistically and economically the quality of the trading strategies devised upon the forecasts.

This study is structured into 6 sections. Section 2 presents a literature review that encompasses several studies on the relationship between cryptocurrencies and conventional financial assets, the transmission of information between cryptocurrencies, and trading strategies. Section 3 presents the raw and transformed data and some descriptive analysis. Section 4 explains the methodologies used to study the information transmission between

Bitcoin and other cryptocurrencies and the approaches to forecast the Bitcoin returns and evaluate trading strategies for this cryptocurrency. Section 5 presents the main results, and the last section concludes the paper.

2. **Literature Review**

Cryptocurrencies have gained attention as both payment methods and investment assets, prompting extensive research on their market dynamics, price determinants, and interactions with traditional financial markets. Studies frequently explore Bitcoin's market efficiency, price drivers, trading volume effects, and its role within the broader financial ecosystem.

Research on Bitcoin's market efficiency shows mixed results. Early studies suggest inefficiency (Kristoufek, 2018; Jiang et al., 2018), but others observe a progression towards efficiency over time (Urquhart, 2016; Wei, 2018). Urquhart (2016) employs randomness tests to find Bitcoin's market-approaching efficiency in recent sub-periods. Wei (2018) expands on this by examining 456 cryptocurrencies and finds a strong relationship between market efficiency, liquidity, and volatility. Conversely, Nadarajah and Chu (2017) conclude Bitcoin is efficient using an alternative methodology.

Balcilar et al. (2017) apply causality-in-quantiles to assess trading volume's impact on Bitcoin's return and volatility, noting predictive power in normal market conditions. This method is advantageous for analysing series with non-Gaussian, asymmetric distributions. Bouri et al. (2019) extend this to seven cryptocurrencies and find volume Granger-causes volatility under low volatility conditions. Dastgir et al. (2019) identify a bidirectional relationship between Bitcoin returns and Google Trends data.

Incorporating cryptocurrencies into the financial market context, Panagiotidis et al. (2018) identify Bitcoin return predictors, including gold returns and internet search intensity. Similarly, Ciner et al. (2022) find significant determinants, including VIX (implied volatility of a hypothetical S&P 500 stock option with 30 days to expiration) and gold prices, during COVID-19. Studies on the interrelation between cryptocurrencies and conventional assets yield conflicting results, with some suggesting market isolation (Ji et al., 2018; Corbet et al., 2018) and others indicating causal connections (Corbet et al., 2020; Bouri et al., 2018). Bitcoin is noted for its safe-haven properties, particularly against equity indices (Shahzad et al., 2019; Corbet et al., 2020).

Research on cryptocurrency interdependence highlights Bitcoin's dominance in information transmission (Koutmos, 2018; Raza et al., 2022), but other studies argue Bitcoin primarily receives information from other cryptocurrencies (Bação et al., 2018; Shahzad et al., 2022). Additionally, studies explore safe-haven and hedge properties (Li et al., 2023; Qiao et al., 2020) and the relationship between cryptocurrencies and fiat currencies (Corelli, 2018; Mokni and Ajmi, 2021). Kim et al. (2021) conclude that there is a significant causal relationship in the tail quantile, which makes it hard for investors to hedge the risk in the cryptocurrency market.

Profitability and trading strategies in cryptocurrency markets are another focus. Manahov (2023) demonstrates consistent profitability despite transaction costs. Momentum effects are explored by Caporale and Plastun (2020) and Bellocca et al. (2022), showing profitable trading patterns. Machine learning models enhance trading profitability (Sebastião and Godinho, 2021; Liu et al., 2023). Other strategies include moving averages (Grobys et al., 2020) and LASSO-based approaches (Huang and Gao, 2022). These studies suggest that machine learning and systematic trading strategies can be effective, robust and profitable.

The mixed evidence on Bitcoin's role in information transmission calls for more robust methodologies to explore interdependencies, particularly using advanced quantile and frequency-based analyses between different time series. Additionally, as machine learning models demonstrate potential in trading strategies, further research should optimize algorithmic approaches to adapt to rapidly changing market conditions and assess their robustness across different market phases.

3. Data and Preliminary Analysis

This study uses daily data retrieved from the CoinMarketCap website (https://coinmarketcap.com/) on the 10 cryptocurrencies with the highest market capitalization on January 1, 2023; excluding stablecoins and cryptocurrencies launched after 2018. These cryptocurrencies, ranked by decreasing market capitalization, are Bitcoin (BTC), Ethereum (ETH), Dogecoin (DOGE), Binance Coin (BNB), Ripple (XRP), Cardano (ADA), Litecoin (LTC), Tron (TRX), Chainlink (LINK), and Ethereum Classic (ETC). The main cryptocurrency under study is BTC and we will refer to other cryptocurrencies as altcoins. Table 1 presents a summary description of these cryptocurrencies on January 1, 2023.

Table 1 – Summary description of cryptocurrencies on January 1, 2023

Crypto	Inception date	Market capitalization USD	Maximum supply	Circulating supply	Price USD	Daily trading volume USD
BTC	Jan. 2009	320,025	21	19	16,625.08	9,244
ETH	Jul. 2015	146,966	n.a.	122	1,200.96	2,400
DOGE	Dec. 2013	132,670	n.a.	132,671	0.070	185
BNB	Jul. 2017	39,053	n.a.	160	244.14	279
XRP	Jun. 2012	17,054	100,000	50,344	0.339	291
ADA	Sep. 2017	8,621	45,000	34,519	0.250	113
LTC	Oct. 2011	5,095	84	726	70.82	344
TRX	Aug. 2017	5,041	n.a.	91,961	0.055	100
LINK	Jun. 2017	2,856	1,000	508	5.622	109
ETC	Jul. 2016	2,188	211	139	15.77	56

Notes: This table presents a summary description of the 10 cryptocurrencies used in this study on January 1, 2023, which are Bitcoin (BTC), Ethereum (ETH), Dogecoin (DOGE), Binance Coin (BNB), Ripple (XRP), Cardano (ADA), Litecoin (LTC), Tron (TRX), Chainlink (LINK), and Ethereum Classic (ETC). The market capitalization, maximum supply, circulating supply, and daily trading volume are presented in millions.

The period under scrutiny is from November 9, 2017, to December 31, 2022 (1,879 daily observations). The raw data includes the closing price, reported at 00:00:00 UTC (Coordinated Universal Time) of the following day, the daily high and low prices, and the daily trading volume in USD. These data were used to compute, for each cryptocurrency , the daily series of logarithmic returns using the closing prices, the log-volumes, the volatility, proxied by the Parkinson range estimator (Parkinson, 1980), and illiquidity, proxied by the Amihud illiquidity ratio (Amihud, 2002).

The Parkinson daily volatility estimator is defined by:

$$\sigma_{i,t}^{p} = \sqrt{\frac{1}{4\ln(2)} ln\left(\frac{H_{i,t}}{L_{i,t}}\right)^{2}}, \tag{1}$$

where $H_{i,t}$ and $L_{i,t}$ are the high and low prices of cryptocurrency i at day t.

Amihud's illiquidity ratio measures the impact on price resulting from a trade of one monetary unit. The daily Amihud illiquidity ratio is defined by:

$$ILLIQ_{i,t} = \frac{|r_{i,t}|}{V_{i,t}}, \tag{2}$$

where $r_{i,t}$ and $V_{i,t}$ correspond to the daily return and trading volume, in USD, of cryptocurrency i at day t. The ratio $ILLIQ_{i,t}$ was then multiplied by 10^8 to have a scale similar to the other variables.

According to the ADF (Augmented Dickey Fuller) test with constant and trend and a number of lags chosen by the BIC (Bayesian Information Criterion), all series are stationary except the log-volumes for some cryptocurrencies. Hence, hereafter we used the first difference of the log-volumes, which are stationary according to the ADF test.

Table 2 shows some descriptive statistics of daily return, first difference of the log-volume, volatility, and illiquidity of the 10 cryptocurrencies. The mean daily returns are very low, with the BNB achieving the highest value of 0.3%. The returns of the cryptocurrencies present a high variability, which is visible by the range and standard deviation. LINK presents the lowest minimum return, -61.5%, while DOGE has the highest maximum return, 151.6%. The standard deviation ranges from 4.0% for BTC to 7.8% for DOGE. Half of the cryptocurrencies have negative skewness and all have excess kurtosis, especially DOGE, with a value of 83.82. Finally, the return series do not show significant first-order autocorrelations, except for ETH and LINK, for the significance levels of 10% and 1%, respectively.

The maximum daily mean first difference of the log-volume is 0.002 (reached by DOGE, BNB, TRX and LINK). The variability is quite high, especially for BNB, with a minimum and a maximum of -9.092 and 9.063, respectively, and a standard deviation of 0.401. However, DOGE shows a higher standard deviation than BNB. All cryptocurrencies present distributions for first difference of the log-volumes with positive skewness, having values ranging from 0.054 for BTC to 1.405 for DOGE, as well as excess kurtosis, with a major highlight of BNB, which presents a value of 277.9. All volume series present significant first-order autocorrelations at a significance level of 1%.

The volatility series proxied by the Parkinson estimator show mean values ranging from 0.003 (BTC) to 0.009 (DOGE and LINK). The maximum value, 1.418, is present in the

DOGE series. As for the standard deviation, DOGE has, once again, the highest value (0.046) and BTC the lowest value (0.006). All series are skewed to the right, with DOGE having the highest value (19.93) and LINK the lowest value (7.655). All series exhibit high excess kurtosis, with DOGE standing out (522.1), and significant first-order autocorrelations at the 1% level.

The illiquidity, proxied by the Amihud ratio, presents average values from 0.000 for BTC to 0.745 for BNB. This last cryptocurrency presents a huge variability of illiquidity, with a minimum very close to 0 and a maximum of 1,335, being much lower in the other cases. As for the standard deviation, BNB stands out again with the highest value (30.83) and BTC has the lowest, very close to zero. All illiquidity series exhibit positive skewness and excess kurtosis, with BNB showing the highest values, 43.30 and 1,873, respectively. There is a significant first-order autocorrelation at a significance of 1% for all series, except BNB.

As expected, BTC stands out as the less volatile cryptocurrency in terms of return, first difference of the log-volumes, volatility, and illiquidity.

Table 2 – Descriptive statistics of returns, volumes, volatility, and illiquidity

	BTC	ETH	DOGE	BNB	XRP	ADA	LTC	TRX	LINK	ETC
Return										
Mean	0.000	0.001	0.002	0.003	0.000	0.001	0.000	0.002	0.002	0.000
Minimum	-0.465	-0.551	-0.515	-0.543	-0.551	-0.504	-0.449	-0.523	-0.615	-0.506
Maximum	0.225	0.235	1.516	0.529	0.607	0.862	0.389	0.787	0.481	0.353
Std dev.	0.040	0.051	0.078	0.059	0.064	0.067	0.055	0.070	0.071	0.063
Skewness	-0.821	-0.908	4.730	0.376	0.828	1.944	-0.129	1.952	-0.078	-0.106
Exc. Kurtosis	11.96	9.561	83.82	14.80	16.40	24.34	8.556	26.01	7.036	7.774
	-0.031	-0.042*	0.017	-0.009	0.004	-0.020	-0.025	0.028	-0.060***	-0.020
Volume										
Mean	0.001	0.001	0.002	0.002	0.000	0.001	0.000	0.002	0.002	-0.000
Minimum	-2.034	-1.083	-1.252	-9.092	-1.963	-1.135	-0.864	-1.029	-4.511	-1.764
Maximum	1.862	1.073	3.981	9.063	2.322	1.801	1.595	2.247	4.438	1.677
Std dev.	0.234	0.235	0.423	0.401	0.355	0.369	0.245	0.282	0.395	0.294
Skewness	0.054	0.296	1.405	0.134	0.682	0.569	0.756	0.984	0.565	0.598
Exc. Kurtosis	6.350	1.636	7.924	277.9	3.480	1.237	3.450	6.108	19.41	3.678
	-0.215***	-0.151***	-0.121***	-0.337***	-0.120***	-0.132***	-0.181***	-0.158***	-0.180***	-0.116***
Volatility										
Mean	0.003	0.004	0.009	0.006	0.007	0.007	0.005	0.008	0.009	0.007
Minimum	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Maximum	0.144	0.192	1.418	0.240	0.316	0.613	0.282	0.571	0.308	0.374
Std dev.	0.006	0.009	0.046	0.016	0.021	0.023	0.012	0.027	0.018	0.018
Skewness	11.45	11.46	19.93	8.850	8.749	15.48	11.09	10.93	7.655	10.55
Exc. Kurtosis	214.9	191.8	522.1	98.75	96.37	338.8	183.8	160.4	86.60	164.2
	0.429***	0.378***	0.337***	0.495***	0.381***	0.426***	0.340***	0.472***	0.455***	0.337***

	BTC	ETH	DOGE	BNB	XRP	ADA	LTC	TRX	LINK	ETC
Illiquidity										
Mean	0.000	0.001	0.111	0.745	0.003	0.040	0.004	0.034	0.619	0.009
Minimum	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Maximum	0.002	0.009	6.213	1,335	0.046	2.470	0.040	6.940	19.58	0.141
Std dev.	0.000	0.001	0.332	30.83	0.005	0.124	0.005	0.290	1.809	0.013
Skewness	3.234	3.418	8.692	43.30	3.670	11.25	2.562	17.19	5.253	3.418
Exc. Kurtosis	12.66	13.75	114.0	1,873	17.13	170.5	7.766	339.8	34.89	17.03
	0.443***	0.448***	0.526***	-0.001	0.527***	0.523***	0.480***	0.536***	0.616***	0.450***

Notes: This table presents the descriptive statistics of daily return, volume (first difference of the logarithm of trading volume), volatility, measured by the Parkinson estimator, and illiquidity, measured by the Amihud illiquidity ratio. These statistics are shown for the 10 cryptocurrencies with the highest market capitalization on January 1, 2023; excluding stablecoins and cryptocurrencies launched after 2018. The cryptocurrencies, ranked by descending market capitalization, are Bitcoin (BTC), Ethereum (ETH), Dogecoin (DOGE), Binance Coin (BNB), Ripple (XRP), Cardano (ADA), Litecoin (LTC), Tron (TRX), Chainlink (LINK), and Ethereum Classic (ETC). The last line for each variable respects the first-order autocorrelation, . Significance at the 10%, 5% and 1% levels are denoted by *, **, ***, respectively. All the series are non-normal at the significance of 1% according to the Jarque-Bera test.

Table 3 presents the correlations between BTC daily returns and one-lagged returns, first difference of the log-volume, volatility, and illiquidity of each cryptocurrency.

Table 3 – Correlations between daily return of BTC and lag return, volume, volatility, and illiquidity of each cryptocurrency

Cryptocurrencies	Returns	Volume	Volatility	Illiquidity
BTC	0.031	0.005	0.058**	0.024
ETH	-0.072***	0.002	0.051*	0.007
DOGE	0.008	0.022	0.028	0.031
BNB	-0.046**	-0.018	0.020	-0.008
XRP	-0.081***	-0.043*	0.031	0.028
ADA	-0.034	0.011	0.042*	0.046**
LTC	-0.063***	0.010	0.037	-0.022
TRX	-0.020	0.022	0.025	0.059**
LINK	-0.023	0.008	0.061***	-0.006
ETC	-0.071***	-0.009	0.050**	0.005

Notes: This table presents the correlations between daily Bitcoin returns and one-lagged return, first difference of the log-volume, volatility, and illiquidity of each of the 10 cryptocurrencies. Significance at the 10%, 5% and 1% levels are denoted by *, **, ***, respectively.

The results presented in Table 3 highlight that the returns of all cryptocurrencies, except DOGE, on day $t - 1$ are negatively correlated with the BTC returns on day t. ETH, XRP, LTC and ETC are significant at the 1% level. The correlations are lower for the other three variables. The only correlation significant at the 1% level is the lagged volatility of LINK, although there are other four cryptocurrencies with volatilities significant at 5% and 10%. The lagged first difference of the log-volume seems to have no information about BTC returns, except for XRP. The lagged illiquidity of ADA and TRX is positively correlated with BTC returns at the 5% significance level, and only the lagged volume of XRP is correlated at the 10% significance level. In a nutshell, this is a clear indication that BTC information is not especially important to forecast its returns, but the inclusion of altcoins in the forecasting models may have significant incremental information.

4. Methodology

This study investigates the Granger causality in the mean and the distribution between BTC and each of the nine most important altcoins in terms of market capitalization. These tests are applied to returns, volume (first difference of the log-volume), volatility and illiquidity. Then, the information on these variables up to time $t - 1$ are used to forecast the value or signal of the BTC return at time t. These signals are then used to devise several trading strategies.

4.1. Granger Causality in the Mean and in the Distribution

The traditional Granger causality test (Granger, 1969) aims to ascertain whether the lags of a potential predictor introduce a significant additional contribution to the prediction of another variable, assuming the linearity of the relationship between the variables. Hence, it tests causality in the mean. Variable X does not Granger causes the variable Y if it does not contribute to its prediction, that is, if:

$$H_0 : f\left(y_t \mid \mathcal{F}_{t-1}^{x\&y}\right) = f\left(y_t \mid \mathcal{F}_{t-1}^{y}\right), \forall Z \epsilon \mathbb{R}, \tag{3}$$

where $f(y_t | \mathcal{F})$ denotes the conditional distribution of y_t, $\mathcal{F}$ the information available at time $t - 1$, such that $\mathcal{F}_{t-1}^{X\&Y}$ corresponds to the information set with the past values of X and Y and $\mathcal{F}_{t-1}^{Y}$ includes only the past values of y , up to time $t - 1$.

To apply this test, it is usual to use bivariate VAR (Vector Autoregressive) models containing only endogenous variables. A VAR model consists of a system of simultaneous equations where each equation presents the contribution of lagged values of the variable itself and other endogenous explanatory variables of the model to the value of the dependent variable, allowing to capture of the linear interdependence relations between the variables. For instance, the equation for variable y_t in a VAR(p) is as follows:

$$y_t = a_0 + \sum_{l=1}^{p} \alpha_l y_{t-l} + \sum_{l=1}^{p} \beta_l x_{t-l} + \varepsilon_t, \tag{4}$$

where a_0 is the constant term, p is the number of lags of stationary variables Y and X, α_l and β_l ($l = 1,...,p$) are the coefficients of the lagged values of Y and X, respectively, and ε_t is the error term.

Variable X does not Granger-causes Y if $H_0 : \beta_1 = ... = \beta_p = 0$. This hypothesis is tested through an F-test, which compares the unrestricted model, including the past values of X and Y, and the restricted model, including only the past values of Y:

$$F = \frac{(SSE_r - SSE_u)/p}{SSE_u / (T - (2p+1))} \quad F_{p,T-(2p+1)} \, if \ H_0 \ is \ true, \tag{5}$$

where SSE_r and SSE_u denote the sum of squared errors from the restricted and unrestricted models, respectively, p is the number of omitted variables in the restricted model and $T - (2p + 1)$ is the number of degrees of freedom, with T corresponding to the number of observations.

The number of lags to include in the VAR was obtained through the multivariate version of the HQC criterion (Hannan-Quinn Criterion) given by $HQC = -2\ell(\hat{\theta}) + 2kloglog(T)$, where $\ell(\hat{\theta})$ is the maximum loglikelihood as a function of the vector of parameter estimates, $\hat{\theta}$ is the vector of estimated parameters, and k is the number of parameters.

The linear causality test causality has been extensively used in macroeconomic and financial applications. More recently, new methodologies have generalized the concept of Granger causality to quantiles and regions of the distribution. The causality test on tail events proposed by Hong et al. (2009) assumes that a tail event occurs when the value of a time series is lower than its VaR (Value-at-Risk) at a specific risk level α%. The $\text{VaR}_{\alpha\%}$ measures the largest possible loss within a confidence interval of α%. The test seeks to determine whether extreme events in a time series contribute to the prediction of extreme events in another time series. The methodology of Hong et al. (2009) has the limitation of being performed in a specific quantile. Candelon and Tokpavi (2016) propose a methodology with a higher testing power, which allows testing Granger causality for several quantiles simultaneously and hence has the flexibility to test specific regions of the distributions supports. Candelon and Tokpavi (2016) is a multivariate extension of Hong et al. (2009), using different VaR levels. For series Y and X:

$$\begin{aligned} Pr\left[Y_t < VaR_t^Y(\theta_Y^0) \mid \mathcal{F}_{t-1}^Y\right] &= \alpha, \\ Pr\left[X_t < VaR_t^X(\theta_X^0) \mid \mathcal{F}_{t-1}^X\right] &= \alpha. \end{aligned} \tag{6}$$

where $VaR_t^Y(\theta_Y^0)$ and $VaR_t^X(\theta_X^0)$ are the VaRs of Y and X, respectively, at time t, and θ_Y^0 and θ_X^0 are the true unknown finite-dimensional parameters related to the VaR models for Y and X, given the information set at time $t-1$.

Let $A = \{\alpha_1,\ldots,\alpha_{m+1}\}$ be a set of $m+1$ VaR levels, covering the distributions support of the variables Y and X, such that $0 < \ldots < \alpha_s \ldots < \alpha_{m+1} < 100\%$, therefore partitioning the support into m disjoint regions. For the series Y, the VaRs at time t are denoted by $VaR_{s,t}^Y(\theta_Y^0,\alpha_s)$, $s = 1,\ldots,m+1$, such that

$$VaR_{1,t}^Y(\theta_Y^0,\alpha_1) < \ldots < VaR_{m+1,t}^Y(\theta_Y^0,\alpha_{m+1}). \tag{7}$$

By convention, $VaR_{s,t}^Y(\theta_Y^0,\alpha_s) = -\infty$ for $\alpha_s = 0\%$ and $VaR_{s,t}^Y(\theta_Y^0\alpha_s) = +\infty$ for $\alpha_s = 100\%$.

The event variable, related to the m disjoint regions of the distribution support of Y, is defined by:

$$Z_{s,t}^Y(\theta_Y^0) = \begin{cases} 1, \textit{ if } Y_t \geq VaR_{s,t}^Y(\theta_Y^0,\alpha_s) \textit{ and } Y_t < VaR_{s+1,t}^Y(\theta_Y^0,\alpha_{s+1}) \\ 0, \textit{otherwise} \end{cases} \tag{8}$$

Note that the event variable for the series X is defined analogously.

Let $H_t^Y(\theta_Y^0)$ and $H_t^X(\theta_X^0)$ be the vectors of dimension $(m, 1)$ that contains the components of the m event variables for series X and Y defined respectively by:

$$\begin{aligned} H_t^Y(\theta_Y^0) &= \left[Z_{1,t}^Y(\theta_Y^0), Z_{2,t}^Y(\theta_Y^0), \ldots, Z_{m,t}^Y(\theta_Y^0)\right]', \\ H_t^X(\theta_X^0) &= \left[Z_{1,t}^X(\theta_X^0), Z_{2,t}^X(\theta_X^0), \ldots, Z_{m,t}^X(\theta_X^0)\right]'. \end{aligned} \tag{9}$$

Then X does not Granger-causes Y in distribution if the following null hypothesis is not rejected:

$$H_0 : \mathbb{E}\left[H_t^Y(\theta_Y^0) \mid \mathcal{F}_{t-1}^{X\&Y}\right] = \mathbb{E}\left[H_t^Y(\theta_Y^0) \mid \mathcal{F}_{t-1}^{Y}\right]. \tag{10}$$

Therefore, Granger causality in distribution from X to Y corresponds to causality in mean for each $H_t^X(\theta_X^0)$ to $H_t^Y(\theta_Y^0)$.

The test can be applied to different regions of the distribution support, such as the centre, the left and right tails, by simply restricting the set $A = \{\alpha_1,...,\alpha_{m+1}\}$ to the desired risk levels. This study considers the left tail by setting $A = \{1\%, 5\%, 10\%\}$, the right tail by setting $A = \{90\%, 95\%, 99\%\}$, and the centre of the distribution, by setting $A = \{20\%, 30\%, 40\%, 50\%, 60\%, 70\%, 80\%\}$.

Let $\hat{H}_t^Y = H_t^Y(\hat{\theta}_Y)$ and $\hat{H}_t^X = H_t^X(\hat{\theta}_X)$ be the estimated counterparts of the multivariate process of the event variables $H_t^Y(\theta_Y^0)$ and $H_t^X(\theta_X^0)$, respectively, with $\hat{\theta}_Y$ and $\hat{\theta}_X$ being the $\sqrt{T}$-consistent estimators of the unknown parameter vectors θ_Y^0 and θ_X^0. $\hat{\Lambda}(j)$ is the sample cross-covariance matrix between $\hat{H}_t^Y$ and $\hat{H}_t^X$ such that:

$$\hat{\Lambda}(j) = \begin{cases} T^{-1} \sum_{t=1+j}^{T} (\hat{H}_t^Y - \hat{\Pi}_Y)(\hat{H}_{t-j}^X - \hat{\Pi}_X)', 0 \leq j \leq T-1, \\ T^{-1} \sum_{t=1-j}^{T} (\hat{H}_{t+j}^Y - \hat{\Pi}_Y)(\hat{H}_t^X - \hat{\Pi}_X)', \quad \textit{otherwise.} \end{cases} \tag{11}$$

The vectors $\hat{\Pi}_Y$ and $\hat{\Pi}_X$ of dimension m are the sample means of $\hat{H}_t^Y$ and $\hat{H}_t^X$, respectively. Like in Hong et al. (2009), $\hat{\Pi}_Y = \mathbb{E}\left[H_t^Y(\theta_Y^0)\right]$ and $\hat{\Pi}_X = \mathbb{E}\left[H_t^X(\theta_X^0)\right]$, without the asymptotic distribution of the test statistic being affected.

The sample cross-correlation matrix, $\hat{R}(j)$, is given by:

$$\hat{R}(j) = D(\hat{\Sigma}_Y)^{-1/2} \hat{\Lambda}(j) D(\hat{\Sigma}_X)^{-1/2}, \tag{12}$$

in which $D(.)$ represents the diagonal form of a matrix and $\hat{\Sigma}_Y$ and $\hat{\Sigma}_X$, which are the sample covariance matrices of $\hat{H}_t^Y$ and $\hat{H}_t^X$, respectively.

Considering further a kernel function $k(.)$, a truncation parameter M and a function $\hat{Q}(j)$, defined by:

$$\hat{Q}(j) = T\left[vec(\hat{R}(j))'\right](\hat{\Gamma}_Y^{-1} \otimes \hat{\Gamma}_X^{-1}) \, vec(\hat{R}(j)), \tag{13}$$

where the operator *vec* vectorises the matrix, $\otimes$ corresponds to the Kronecker product, and $\hat{\Gamma}_Y$ and $\hat{\Gamma}_X$ are the sample correlation matrices of $\hat{H}_t^Y$ and $\hat{H}_t^X$, respectively.

The test statistic associated with the null hypothesis of non-causality can be represented by a weighted quadratic form that considers the dependence between the current value of $\hat{H}_t^Y$ and the lagged values of $\hat{H}_t^X$, that is, by:

$$\hat{\mathfrak{I}} = \sum_{j=1}^{T-1} k^2\left(\frac{j}{M}\right)\hat{Q}(j). \tag{14}$$

The test statistic of Candelon and Tokpavi (2016) is a centred and scaled version of the quadratic form present in the previous equation:

$$V_{X\to Y} = \frac{\hat{\mathfrak{I}} - m^2 C_T(M)}{\left(m^2 D_T(M)\right)^{1/2}}, \tag{15}$$

where $C_T(M)$ and $D_T(M)$ are the location and scale parameters, corresponding respectively to:

$$C_T(M) = \sum_{j=1}^{T-1}(1 - j/T)\,k^2(j/M), \tag{16}$$

$$D_T(M) = 2\sum_{j=1}^{T-1}(1 - j/T)(1 - (j+1)/T)\,k^4(j/M). \tag{17}$$

Under the null hypothesis of no causality in distribution, $V_{X\to Y} \sim N(0,1)$.

As discussed by Hong et al. (2009), the choice of kernel, except for the case of the uniform kernel that does not eliminate higher-order lags, is not relevant as it leads to comparable test powers. In this study, we will resort to the Bartlett kernel.

Candelon and Tokpavi (2016) consider three values for the truncation parameter M, namely $ln(T)$, $1.5T^{0.3}$, and $2T^{0.3}$. We have tested these values with similar results. Hence, results are presented for $1.5T^{0.3}$, which, given the sample size, is 14.

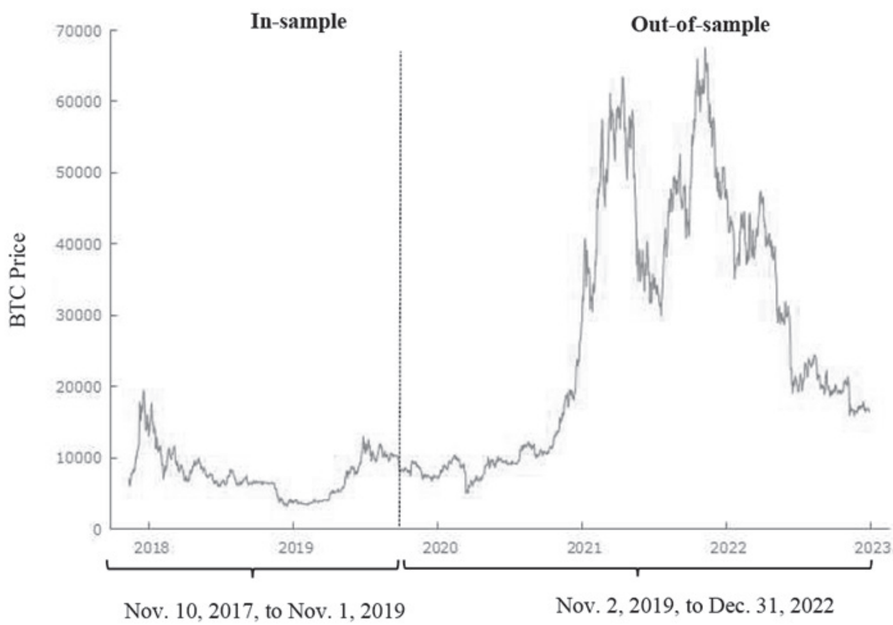
4.2. Forecasting and Trading on Bitcoin

This subsection explains the procedures used to forecast BTC returns and to devise trading strategies based on those forecasts.

The total period was partitioned into in-sample and out-of-sample. The in-sample is from November 10, 2017, to November 1, 2019, and the out-of-sample period is from November 2, 2019, to December 31, 2022, so that $T_1 = 722$ and $T_2 = 1{,}156$ (see Figure 1). A rolling window with a fixed length of 714 observations, was used to forecast BTC returns based on the lagged information on returns, volumes, volatility, and illiquidity series of BTC and the other nine altcoins. We consider 11 models with BTC returns as the dependent variable and different predictor spaces with lags of 1 to 7 to capture any day-of-the-week effect.

One model only considers BTC, nine models use BTC and an altcoin, and the last model uses all the information.

Figure 1 – Evolution of Bitcoin closing price and partition into in- and out-of-sample.



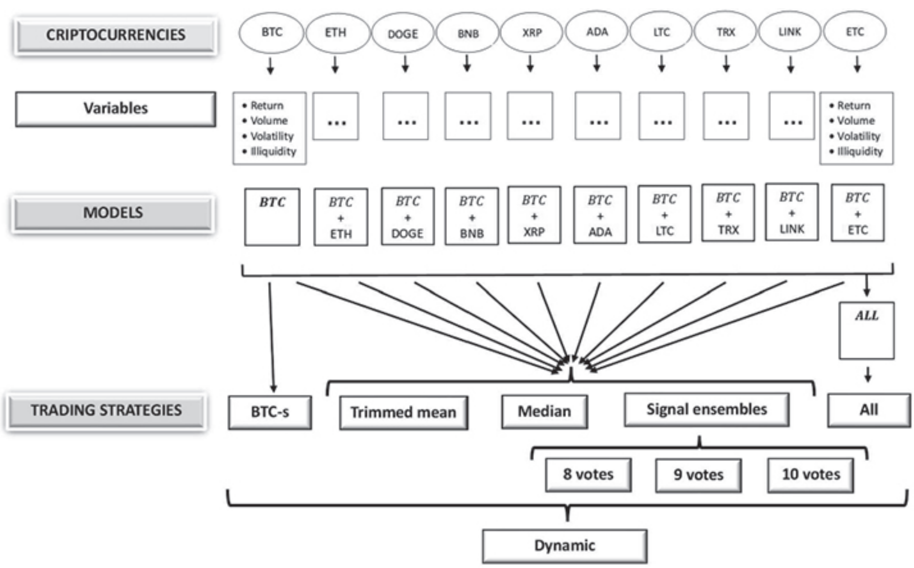
The forecasting models are then used to devise the following eight trading strategies:

- Strategy "BTC-s" considers the information of BTC only (7 x 4 = 28) explanatory variables.
- Strategies "Trimmed mean" and "Median" use the trimmed mean and median of the forecasts obtained from ten regressions, one for each cryptocurrency with 28 explanatory variables, respectively. The trimmed mean is the arithmetic average of the six forecasts, excluding the two lowest and two highest forecasts.
- The "Signal ensembles", consider the signals of the ten forecasts obtained from the model with BTC and from the nine models with the information of BTC and each cryptocurrency (in these models, there are 2 x 7 x 4 = 56 explanatory variables). The investor enters, stays in, stays out or exits the market if eight, nine and ten models agree on the signal of the forecasts (these strategies are called hereafter "8 votes", "9 votes" and "10 votes", respectively).
- Strategy "All" considers as the predictor space all the lagged information of the ten cryptocurrencies (in total 10 x 7 x 4 = 280 variables).

- Finally, the "Dynamic" strategy chooses the best one, step-by-step, among the previous seven strategies. For each day, the best strategy is the one that minimizes the MSE (mean squared error) of the previous seven days (in case of a tie, the strategy with the model with the highest cumulative return in the previous days is adopted).

Figure 2 illustrates the research framework for trading strategies, referring to the cryptocurrencies, variables, forecasting models, and trading strategies used.

Figure 2 – Research framework for trading strategies



The models were estimated using the LASSO (Least Absolute Shrinkage and Selection Operator) procedure proposed by Tibshirani (1996) to remove redundant variables and select the relevant regressors. The LASSO estimator is defined as:

$$\hat{\beta}_{LASSO} = \operatorname{argmin}\left\{\sum_{t=8+k}^{t_1+k}\left(y_t - \beta_0 - \sum_j \beta_j x_{tj}\right)^2 + \lambda\sum_j \beta_j^2\right\}, \lambda \geq 0, \tag{18}$$

where y_t and x_{tj} correspond to the observations of the dependent and the explanatory variables j at time t, respectively, out of a total of $T_1 - 7$ observations, where T_1 is the in-sample size, and the optimization problem is solved for $k = 0,\ldots,T_2 - 1$, where T_2 is the end of the out-of-sample. β_j is the regression coefficient corresponding to x_j and β_0 is the constant term of the model. Finally, λ is the regularisation parameter (penalty), which allows the elimination of redundant coefficients. The higher its value, the higher the penalty and the number of null coefficients.

The choice of the regularization parameter was carried out through a 10-fold cross-validation. Briefly, this method divides the period into ten disjoint subsets of approximately equal size and trains the model in nine subsets which are applied to the remaining subset. This procedure is replicated ten times. The cross-validation performance corresponds to the average of a performance measure across the ten subsets. In our case, the λ is the one that minimizes the MSE in the previous week, that is:

$$MSE = \frac{1}{T_1 - 7} \sum_{t=8+k}^{T_1+k} (y_t - \hat{y}_t)^2, \tag{19}$$

where y_t and $\hat{y}_t$ are the observed and estimated value of the dependent variable, respectively, and $k = 0,...,T_2 - 1$.

The trading strategies only consider positive or null positions in BTC, implying that short selling is precluded. Hence the investor is unable to capitalize on negative forecasts. The action taken by the investor depends on the forecasts of the trading strategies and a threshold. This threshold is set equal to the proportional transaction costs of 0.25% which is higher than the figures used in the literature for BTC (Alessandretti et al., 2018, Sebastião and Godinho, 2021). The procedure is the following. If the forecast at day t is higher than 0.25%, the investor enters or stays in the market at day $t + 1$ if at day t the position is null or positive, respectively. If the forecast at day t is lower than -0.25%, the investor exists or stays out of the market at day $t + 1$ if at day t the position is positive or null, respectively. The Signal ensembles add an additional step. For each model (BTC and BTC and each altcoin) the forecast is made, and the signal is recorded (1 if the forecast is higher than 0.25% and 0 if the forecast is lower than -0.25%). If the number of models with a given signal is equal to or higher than a given boundary the investor takes action. For instance, for the strategy "8 votes" the investor gets out or stays out of the market if the sum of ones is lower than 8 and enters or stays in the market otherwise.

The three best strategies according to the Sharpe ratio are compared with the passive Buy-and-hold (B&H) strategy and the BTC-s strategy. To compare the performance of all strategies, the strategies are analysed from $T_1 + 8$ to T_2, 1,149 observations (hence excluding the first 7 observations of the out-of-sample period, due to the dynamic strategy). Several performance metrics are computed with proportional trading costs and entry/exit barrier of 0.25%.

(1) The relative number of days in the market in which a long position is active.

(2) The win rate corresponds to the percentage of days in the market in which the strategy returns a positive return.

(3) The cumulative return after trading costs given by the exponential of the sum of the daily continuous returns of strategy over the entire evaluation, i.e., $\exp(\sum r_{j,t}) - 1$.

(4) The annualized mean return.

(5) The annualized standard deviation of returns.

(6) Assuming that the risk-free rate is zero, the annualized Sharpe ratio is the ratio between the return, $\hat{\mu}_j$, and the standard deviation, $\hat{\sigma}_j$:

$$SR_j = \sqrt{365}\frac{\hat{\mu}_j}{\hat{\sigma}_j}. \tag{20}$$

(7) The bootstrap p-values corresponding to the probabilities of the daily Sharpe ratio of the active strategy, considering all days in the sample, are higher than the daily Sharpe ratio of the B&H and BTC-s strategies.

(8) The annualized Sortino ratio which considers in the denominator the downside risk from a target value, which we assume is equal to zero:

$$STR_j = \sqrt{365}\frac{\hat{\mu}_j}{\sqrt{\frac{1}{T_2 - T_1 - 8}\sum_{t=T_1+8}^{T_2} \min\left[r_{j,t}, 0\right]^2}} \tag{21}$$

(9) The annualized certainty equivalent of a CRRA (Constant Relative Risk Aversion) utility function such that:

$$U(W_t) = \begin{cases} \frac{W_t^{1-\gamma}}{1-\gamma}, & \text{if } \gamma > 1 \\ \ln(W_t), & \text{if } \gamma = 1 \end{cases} \tag{22}$$

where W_t denotes investor wealth at t and γ is the risk aversion parameter is given by:

$$CE_j = \begin{cases} \left[\left(\frac{1}{T_2 - T_1 - 8}\sum_{t=T_1+8}^{T_2} \hat{W}_{j,t}^{1-\gamma}\right)^{\frac{365}{1-\gamma}} - 1\right] & \text{if } \gamma \neq 1 \\ 365\left[\frac{1}{T_2 - T_1 - 8}\sum_{t=T_1+8}^{T_2} \log \hat{W}_{j,t}\right] & \text{if } \gamma = 1 \end{cases}, \tag{23}$$

where $\hat{W}_t = e^{r_{j,t}}$ (we considered $\gamma = 1, 3, 5$).

(10) Lastly, the $\text{CVaR}_{\alpha\%}$ (Conditional Value-at-Risk at α%) measures the average loss conditional on a VaR exceeded at the $\alpha\% = 1\%, 5\%$.

5. Results

5.1. Causality in the Mean and in the Distribution

Table 4 and Table 5 present the tests on Granger causality in the mean and in the distribution, respectively, between BTC and the altcoins. In discussing these results, we mainly focus on those that are significant at the 1% level.

Table 4 – Granger causality in the mean between Bitcoin and each altcoin

	Returns		Volume		Volatility		Illiquidity	
i	$F_{BTC \to i}$	$F_{i \to BTC}$	$F_{BTC \to i}$	$F_{i \to BTC}$	$F_{BTC \to i}$	$F_{i \to BTC}$	$F_{BTC \to i}$	$F_{i \to BTC}$
ETH	1.140	5.236**	2.699***	1.498	2.339*	3.052**	2.081**	1.602
DOGE	4.208***	0.745	2.883***	1.738*	1.168	1.791	1.559	1.494
BNB	2.137	1.591	5.229***	2.103**	3.020**	3.101***	0.222	3.893***
XRP	3.807*	6.138**	3.803***	1.489	0.683	1.709	3.520***	1.498
ADA	1.238	4.146**	5.270***	2.209**	2.065**	2.459**	2.556**	2.703***
LTC	0.405	3.732*	2.635**	0.665	8.030***	7.644***	0.706	2.704***
TRX	2.383*	0.037	3.629***	1.750*	2.805***	2.559**	1.524	3.337***
LINK	1.031	0.045	3.443***	2.482**	4.502***	4.403***	0.633	1.668
ETC	1.480	4.530**	1.906*	2.079**	2.344*	1.953*	1.739*	2.375**

Notes: $F_{BTC \to i}$ and $F_{i \to BTC}$ denote the statistics of the linear Granger causality test from BTC to altcoin i and from altcoin i to BTC, respectively. Significance at the 10%, 5% and 1% levels are denoted by *, **, ***, respectively.

Granger causality in the mean runs only from the returns of BTC to DOGE. In terms of the first difference of the log-volume, the causality runs from BTC to all altcoins, except LTC and ETC. At the 1% significance level, there is bidirectional causality between BTC and LTC and BTC and LINK. Additionally, at this significance level, there is causality from BTC to TRX and from BNB to BTC. The causality in illiquidity runs mainly from altcoins to BTC, namely from BNB, ADA, LTC and TRX. Only BTC Granger causes the XRP illiquidity at the 1% level.

Table 5 shows the Granger causality tests in distribution, applied to the left tail (bearish market), right tail (bull market) and the central region (calm market) between BTC and the nine altcoins for return, volume, volatility, and illiquidity.

Table 5 – Granger causality in the distribution between Bitcoin and each altcoin

	Left tail		Centre		Right tail	
	$V_{BTC \to i}$	$V_{i \to BTC}$	$V_{BTC \to i}$	$V_{i \to BTC}$	$V_{BTC \to i}$	$V_{i \to BTC}$
Return						
ETH	-0.031	0.497	0.622	0.130	0.226	-1.006
DOGE	0.260	0.406	-0.253	0.716	-0.494	0.948
BNB	3.163***	0.750	-0.987	3.243***	-1.333	-0.850
XRP	1.133	1.541	0.297	1.020	0.915	2.148**
ADA	0.229	-1.153	1.250	-0.217	-1.223	-0.934
LTC	2.316**	-0.673	0.116	1.038	0.813	-0.392

	Left tail		Centre		Right tail	
	$V_{BTC\to i}$	$V_{i\to BTC}$	$V_{BTC\to i}$	$V_{i\to BTC}$	$V_{BTC\to i}$	$V_{i\to BTC}$
Return						
TRX	1.927*	1.133	-0.994	2.959***	-0.309	1.470
LINK	0.872	-1.004	-1.098	2.204**	-0.879	-1.457
ETC	2.350**	-0.659	0.232	1.097	-0.119	-0.072
Volume						
ETH	4.234***	4.787***	-0.262	-0.215	-0.586	0.604
DOGE	1.898*	2.105**	3.038***	-0.936	0.547	0.199
BNB	2.708***	-0.106	-0.354	-0.655	1.485	0.408
XRP	5.521***	4.293***	-1.309	-1.576	0.088	-0.050
ADA	3.688***	0.451	-0.937	1.347	0.784	1.096
LTC	2.365**	1.309	1.864*	-0.389	0.617	-1.970**
TRX	1.683*	0.063	-1.360	-0.734	2.071**	-1.798*
LINK	1.521	1.119	-0.830	-0.597	-1.295	-0.328
ETC	0.251	1.579	0.4704	0.162	-1.425	-1.129
Volatility						
ETH	11.16***	9.063***	0.497	-0.484	2.662***	-0.834
DOGE	1.384	-0.581	1.702*	-0.246	1.060	2.085**
BNB	5.927***	2.3008**	-0.226	2.849***	1.131	0.858
XRP	4.293***	1.435	1.188	0.432	0.241	-0.521
ADA	3.794***	2.137**	-0.831*	1.497	0.225	0.589
LTC	9.650***	1.512	0.427	0.692	1.211	-0.406
TRX	5.268***	2.942***	0.434	-1.260	3.152***	1.566
LINK	2.016**	1.627	0.297	1.127	0.739	-0.356
ETC	6.391***	5.784***	0.858	0.032	1.186	1.591
Illiquidity						
ETH	1.050	-0.930	-0.411	-0.228	-0.314	1.752*
DOGE	-1.769*	1.465	1.485	1.353	-0.972	-0.341
BNB	2.543***	0.330	0.034	-0.293	6.689***	2.614***
XRP	0.225	-0.477	-1.211	-0.367	-2.213**	0.534
ADA	-1.211	-0.369	1.121	0.055	0.213	1.117
LTC	-1.238	0.027	-1.582	-2.103**	-0.561	1.431
TRX	-1.431	-1.268	0.734	-1.086	0.481	4.790***
LINK	0.179	-1.692*	-0.441	0.112	-1.663*	-1.831*
ETC	-0.431	1.871*	-2.162**	-0.334	-1.169	-0.698

Notes: This table shows the causality test in the distribution of Candelon and Tokpavi (2016) applied to the left tail (quantiles 0.01, 0.05 and 0.1), right tail (quantiles 0.9, 0.95 and 0.99), and centre of the distribution (quantiles 0.2, 0.3, 0.4, 0.5, 0.6, 0.7 and 0.8). $V_{BTC \to i}$ and $V_{i \to BTC}$ denote the causality statistics from BTC to the altcoins and vice-versa, respectively. The tests were performed using the Bartlett kernel and a truncation parameter $M = 1.5T^{0.3} = 14$. Significance at the 10%, 5% and 1% levels are denoted by *, **, ***, respectively.

For returns, in the left tail, the causality runs from BTC to four altcoins (BNB, LTC, TRX and ETC) but is only significant at the 1% level for BNB. In the centre, causality runs from BNB to BTC, while in the right tail, there is no significant causality at the 1% level. For first difference of the log-volume, most of the causality occurs in the left tail where there are two cases of bidirectional causality (ETH and XRP) at the 1% level. Most of these causalities fade away in the centre and especially in the right tail. The exception is DOGE, in the centre of the distribution, where now the causality from BTC to DOGE is reinforced. Volatility presents a similar pattern but with more positive results. In the left tail, most of causality runs from BTC to the altcoins, with DOGE being the only altcoin without any significant causality. There is bidirectional causality between BTC and ETH, TRX, and ETC. In the presence of bullish markets, there is causality from BTC to ETH and TRX. Illiquidity presents a scarcer number of significant relationships. In the left tail BTC causes BNB, in the centre there is no significant relationship at 1%, and, interestingly, in the right tail, the causality runs bidirectionally between BTC and BNB, and unidirectionally from TRX to BTC.

5.2. Performance of the Trading Strategies

The three trading strategies with the highest Sharpe ratio at the end of the out-of-sample period are the voting system "9 votes", the strategy based on the lagged information of all cryptocurrencies ("All") and the "Dynamic" system. These three strategies are assessed out-of-sample, and the results of their performance are presented in Table 6. This table also presents the results of the strategy based only on BTC information. Clearly, this is a poor strategy providing almost the same results as the B&H strategy.

Although there are mixed results in terms of win rate, certainty equivalent and extreme risk, measured by the VaR, we may claim that the three best strategies outperform the B&H strategy, and the strategy based only on BTC information. Most notably, "All" is the one with the best results in the most important metrics, i.e. Sharpe ratio and Sortino ratio. In all dimensions analysed the "All" strategy beats by far B&H strategy. The "All" strategy provides a cumulative return after transaction costs of 331.4%, while the B&H strategy, with no transaction costs, has a cumulative return of 187.9%. The higher mean return coupled with the lower standard deviation of the "All" strategy provides a Sharpe ratio of 94.59%, which is higher than the Sharpe ratio of the B&H strategy at the 10% significance level. The claim on the superiority of the "All" strategy is reinforced by the Sortino ratio, which achieves a value of 139.7%. The strategy is better suited for investors with low-risk aversion ($\gamma = 1$), although in terms of extreme risk, measured by the VaR, is comparable to the other two best strategies.

Table 6 – Performance of the best three trading strategies after round-trip transaction costs of 0.5% and an entry/exit barrier of ±0.25%

			Best 3 strategies		
	B&H	BTC-s	9 votes	All	Dynamic
Percentage of days in the market	100	83.72	45.43	60.14	63.19
Win rate	51.00	51.35	53.64	52.68	52.48
Cumulative return	187.9	175.0	221.9	331.4	281.1
Annualized mean return	46.31	41.36	38.00	56.40	49.20
Annualized std. deviation	71.54	67.69	50.44	59.63	57.40
Annualized Sharpe ratio	64.74	61.11	75.34	94.59	85.71
Bootstrap p-values against B&H	--	50.19	24.56	9.99	17.98
Bootstrap p-values against BTC	50.10	--	23.56	9.98	17.53
Annualized Sortino ratio	92.01	86.57	113.5	139.7	131.0
Annualized CE with $\gamma = 1$	20.04	17.77	25.31	38.07	32.83
Annualized CE with $\gamma = 3$	-30.39	-28.22	-0.15	-1.79	0.16
Annualized CE with $\gamma = 5$	-63.88	-60.72	-22.93	-39.57	-27.95
CVaR at 1%	14.33	14.30	10.67	12.39	10.98
CVaR at 5%	8.60	8.38	6.59	7.09	7.11

Notes: This table presents the performance of the three best strategies, according to the Sharpe ratio, and compares them with the Buy-and-Hold (B&H) and the active strategy that only uses BTC information (BTC-s). The best active strategies are the Signal ensemble with 9 votes (denoted by "9 votes"), the strategy based on the model forecasts with all information of the 10 cryptocurrencies (denoted by "All") and the strategy that chooses dynamically the best strategy out of the 8 active strategies considered. Besides the relative number of days with an active long position in the market, the strategies are assessed with the following performance metrics: Win rate corresponding to the percentage of days in the market in which the strategy has a positive return, Cumulative return given by the exponential of the of the daily continuous returns of strategy j, $\exp(\Sigma\, r_{j,t}) - 1$, annualized mean, annualized standard deviation, annualized Sharpe ratio, assuming that the risk-free rate is zero, bootstrap p-values corresponding to the probabilities of the daily Sharpe ratio of the active strategy are higher than the daily Sharpe ratio of the B&H and of the BTC-s strategies, annualized Sortino ratio which considers in the denominator the downside risk from a target value equal to zero, annualized certainty equivalent of a CRRA (Constant Relative Risk Aversion) utility function with a risk aversion parameter of $\gamma = 1, 3, 5$, and the $CVaR_{\alpha\%}$ (Conditional Value-at-Risk at $\alpha\%$) with $\alpha\% = 1\%, 5\%$. All metrics are computed on returns after round-trip transaction costs of 0.5%. The p-values were obtained using 100,000 bootstrap samples created with the circular block procedure of Politis and Romano (1994), with an optimal block size chosen according to Politis and White (2004) and Patton et al. (2009). All values are in percentage.

Although the strategies are assessed considering a threshold of ±0.25%, which is an obvious figure due to the consideration of round-trip transaction costs of 0.5%, arguably the profitability of the trading strategies could be fostered by optimizing this parameter. Table 7 presents a sensitivity analysis, considering several entry and exit barriers.

Table 7 – Sensitivity of the trading strategies to the market enter and exit thresholds

Thresholds	Statistics	BTC-s	Best of other strategies	All	Dynamic
±0.2%	Rank (#)	(#5)	9 votes (#3)	(#1)	(#2)
	SR (%)	49.83	53.91	84.14	76.79
	p-value	0.753	0.452	0.168	0.263
	Σrt (%)	137.2	158.3	271.9	241.0
±0.25%	Rank (#)	(#7)	9 votes (#3)	(#1)	(#2)
	SR (%)	61.11	75.34	94.59	85.71
	p-value	0.502	0.246	0.099*	0.180
	Σrt (%)	175.0	221.9	331.4	281.0
±0.3%	Rank (#)	(#8)	Median (#2)	(#1)	(#5)
	SR (%)	60.19	102.5	104.6	80.55
	p-value	0.510	0.071*	0.053*	0.220
	Σrt (%)	171.6	399.3	402.0	253.4
±0.35%	Rank (#)	(#8)	10 votes (#2)	**(#1)**	(#7)
	SR (%)	68.75	116.28	**118.28**	70.71
	p-value	0.334	0.046**	**0.016****	0.315
	Σrt (%)	205.7	341.6	**533.7**	214.5
±0.4%	Rank (#)	(#5)	10 votes (#2)	(#1)	(#8)
	SR (%)	91.48	112.2	113.4	60.56
	p-value	0.124	0.054*	0.022**	0.417
	Σrt (%)	328.4	307.7	488.9	178.1
±0.45%	Rank (#)	(#6)	Median (#1)	(#2)	(#8)
	SR (%)	91.01	109.3	107.8	39.31
	p-value	0.127	0.043**	0.037**	0.693
	Σrt (%)	326.4	453.7	434.8	116.1
±0.5%	Rank (#)	(#5)	Median (#2)	(#1)	(#8)
	SR (%)	97.72	102.5	115.3	48.97
	p-value	0.080*	0.072*	0.022**	0.583
	Σrt (%)	373.2	390.3	499.6	139.4
±0.55%	Rank (#)	(#5)	Trim. mean (#2)	(#1)	(#8)
	SR (%)	78.23	91.52	110.39	17.74
	p-value	0.252	0.118	0.036**	0.879
	Σrt (%)	250.6	329.5	445.0	77.46

Notes: This table presents a sensitivity analysis of the trading strategies to the market entry and exit thresholds. Rank refers to the order of the strategy out of the overall 8 strategies (BTC-s, Trimmed mean, Median, 8 votes, 9 votes, 10 votes, All, and Dynamic) according to the Sharpe ratio. SR is the Sharpe ratio, p-value is the bootstrap p-value against

B&H, i.e., the probability of the daily Sharpe ratio of the active strategy, considering all days in the sample, being higher than the Sharpe ratio of B&H strategy that consists of being long all the time (these p-values are obtained using 100,000 bootstrap samples created with the circular block procedure of Politis and Romano (1994), with an optimal block size chosen according to Politis and White (2004) and Patton et al. (2009)) and Σr_t is the cumulative return of the strategy out-of-sample. Significance at the 10%, 5% and 1% levels are denoted by *, **, ***, respectively. The best threshold and best strategy with that threshold are highlighted in bold.

The results point out that the "Dynamic" strategy is highly sensitive to the threshold. With lower thresholds, this is the second-best strategy, but when the threshold increases the performance of this strategy decreases and at thresholds higher than ±0.4% it is ranked as the worst strategy of all. The "All" strategy is always the best strategy, except when the threshold is equal to ±0.45%, however, the difference in the Sharpe ratios between the "All" strategy and the best strategy is only 1.5%. Finally, it seems that the performance of the strategies is a concave function of the threshold value. For the "All" strategy and the best other strategy, the best performance is achieved at an entry/exit barrier of ±0.35%.

6. Conclusion

This paper investigates the information transmission concerning the returns, volumes, volatilities and illiquidity between BTC and the nine altcoins with the highest market capitalization, excluding stable cryptocurrencies and those created after 2018, between November 10, 2017, and December 31, 2022. This was accomplished using causality tests in the mean and the distribution.

Contrary to the claim of several studies (e.g., Koutmos, 2018; Ji et al., 2019), there is no clear dominance of BTC regarding information transmission to altcoins in the mean. In terms of returns and illiquidity, most causality runs in the opposite direction, from altcoins to BTC, which is in line with Bação et al. (2018). In terms of volatility, the causality is mainly bidirectional as claimed by Ji et al. (2019) and Raza et al. (2022). However, BTC shows its superiority in terms of volume.

The causal relationship between BTC and each altcoin is more evident in the left tail of the distributions, except for illiquidity where the right tail stands out, with the transmission of information occurring mainly from BTC to altcoins. On the other hand, at the highest quantiles, causality occurs mainly from altcoins to BTC (except for volume). This agrees with the results obtained by Shahzad et al. (2022).

These results suggest that there is some information transmission from altcoins to BTC, especially at the highest quantiles of the distribution, and this information can be used profitably to trade in BTC. To assess this hypothesis, one used the following procedure: (1) Eleven models were built, with the first one having the BTC series as explanatory variables, nine using the information from BTC and each one of the altcoins considered, and a last one with all series, with a predictor space formed by 280 variables. (2) The forecasts were obtained dynamically day by day using a moving window with a fixed length. Given the high dimensionality of the optimization problem, we resort to LASSO regressions, which allow the selection of the important regressors. (3) Finally, the performance of the trading strategies which use a combination of the forecasts and all the information on the

10 cryptocurrencies was assessed out-of-sample using several metrics and considering round-trip transaction costs of 0.5% and an entry/exit barrier of ±0.25%.

The strategy based only on BTC information has very low performance providing almost the same results as the B&H strategy. Hence, we may conclude that if there is some information in the BTC, this information is not important enough to surpass the transaction costs. The three trading strategies with the highest Sharpe ratio at the end of the out-of-sample period are the voting system "9 votes", according to which the investor enters, stays in, or exits the market if nine of the ten models agree on the signal of the forecasts, the strategy based on the lagged information of all cryptocurrencies and the "Dynamic" system, which chooses step-by-step the strategy that minimizes the MSE (mean squared error) of the previous seven days. The best strategy is the one that uses the information of all cryptocurrencies to forecast, via LASSO, the Bitcoin returns. This strategy provides a cumulative return after transaction costs of 331.4%, while the B&H strategy, with no transaction costs, has a cumulative return of 187.9%, a Sharpe ratio of 94.59%, which is higher than the Sharpe ratio of the B&H strategy at the 10% significance level. The strategy is better suited for investors with low-risk aversion ($\gamma = 1$) with an extreme risk lower than the B&H strategy.

Finally, we tested the sensitivity of the performances to the trading rules on the entry/exit threshold. The strategy with all information presents robust results in the face of varying thresholds, being almost always the best strategy. It seems that the performance of the strategies is a concave function of the threshold value, and hence the profitability of the trading strategies may be fostered by optimizing this parameter.

All in all, probably the most important conclusion to retrieve from this paper is that trading strategies on Bitcoin should consider large sets of predictors, namely the information from other cryptocurrencies. This claim is of outmost important for investors in cryptocurrencies.

Based on the results presented here, future work may be developed considering an expanded sample of cryptocurrencies, using machine learning models, and optimizing the parameter of transaction trigger.

References

Alessandretti, L.; ElBahrawy, A.; Aiello, L. M.; Baronchelli, A. (2018) Anticipating cryptocurrency prices using Machine Learning. *Complexity*, 2018, 1-16. https://doi.org/10.1155/2018/8983590

Amihud, Y. (2002) Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1), 31-56. https://doi.org/10.1016/s1386-4181(01)00024-6

Bação, P.; Duarte, A. P.; Sebastião, H.; Redzepagic, S. (2018) Information transmission between cryptocurrencies: Does Bitcoin rule the cryptocurrency world? *Scientific Annals of Economics and Business*, 65(2), 97-117. https://doi.org/10.2478/saeb-2018-0013

Balcilar, M.; Bouri, E.; Gupta, R.; Roubaud, D. (2017) Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling*, 64, 74-81. https://doi.org/10.1016/j.econmod.2017.03.019

Bellocca, G. P.; Attanasio, G.; Cagliero, L.; Fior, J. (2022) Leveraging the momentum effect in machine learning-based cryptocurrency trading. *Machine Learning with Applications*, 8, 100310. https://doi.org/10.1016/j.mlwa.2022.100310

Bouri, E.; Gupta, R.; Lau, C. K. M.; Roubaud, D.; Wang, S. (2018) Bitcoin and global financial stress: A copula-based approach to dependence and causality in the quantiles. *The Quarterly Review of Economics and Finance*, 69, 297-307. https://doi.org/10.1016/j.qref.2018.04.003

Bouri, E.; Lau, C. K. M.; Lucey, B.; Roubaud, D. (2019) Trading volume and the predictability of return and volatility in the cryptocurrency market. *Finance Research Letters*, 29, 340-346. https://doi.org/10.1016/j.frl.2018.08.015

Candelon, B.; Tokpavi, S. (2016) A Nonparametric test for Granger causality in distribution with application to financial contagion. *Journal of Business & Economic Statistics*, 34(2), 240-253. https://doi.org/10.1080/07350015.2015.1026774

Caporale, G. M.; Plastun, A. (2020) Momentum effects in the cryptocurrency market after one-day abnormal returns. *Financial Markets and Portfolio Management*, 34(3), 251-266. https://doi.org/10.1007/s11408-020-00357-1

Ciner, C.; Lucey, B.; Yarovaya, L. (2022) Determinants of cryptocurrency returns: A LASSO quantile regression approach. *Finance Research Letters*, 49, 102990. https://doi.org/10.1016/j.frl.2022.102990

Comissão do Mercado de Valores Mobiliários. (2022) O que são criptoativos? Retrieved on October 2, 2022, from https://www.cmvm.pt/pt/AreadoInvestidor/Faq/Pages/FAQs-Criptoativos_investidores.aspx

Corbet, S.; Katsiampa, P.; Lau, C. K. M. (2020) Measuring quantile dependence and testing directional predictability between Bitcoin, altcoins and traditional financial assets. *International Review of Financial Analysis*, 71, 101571. https://doi.org/10.1016/j.irfa.2020.101571

Corbet, S.; Meegan, A.; Larkin, C.; Lucey, B.; Yarovaya, L. (2018) Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Economics Letters*, 165, 28–34. https://doi.org/10.1016/j.econlet.2018.01.004

Corelli, A. (2018). Cryptocurrencies and exchange rates: A relationship and causality analysis. *Risks*, 6(4), 111. https://doi.org/10.3390/risks6040111

Dastgir, S.; Demir, E.; Downing, G.; Gozgor, G.; Lau, C. K. M. (2019) The causal relationship between Bitcoin attention and Bitcoin returns: Evidence from the copula-based Granger causality test. *Finance Research Letters*, 28, 160-164. https://doi.org/10.1016/j.frl.2018.04.019

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 424-438. https://doi.org/10.2307/1912791

Grobys, K.; Ahmed, S.; Sapkota, N. (2020) Technical trading rules in the cryptocurrency market. *Finance Research Letters*, 32, 101396. https://doi.org/10.1016/j.frl.2019.101396

Hong, Y.; Liu, Y.; Wang, S. (2009) Granger causality in risk and detection of extreme risk spillover between financial markets. *Journal of Econometrics*, 150(2), 271-287. https://doi.org/10.1016/j.jeconom.2008.12.013

Huang, W.; Gao, X. (2022) LASSO-based high-frequency return predictors for profitable Bitcoin investment. *Applied Economics Letters*, 29(12), 1079-1083. https://doi.org/10.1080/13504851.2021.1908512

Ji, Q.; Bouri, E.; Gupta, R.; Roubaud, D. (2018) Network causality structures among Bitcoin and other financial assets: A directed acyclic graph approach. *The Quarterly Review of Economics and Finance*, 70, 203-213. https://doi.org/10.1016/j.qref.2018.05.016

Ji, Q.; Bouri, E.; Lau, C. K.; Roubaud, D. (2019) Dynamic connectedness and integration in cryptocurrency markets. *International Review of Financial Analysis*, 63, 257-272. https://doi.org/10.1016/j.irfa.2018.12.002

Jiang, Y.; Nie, H.; Ruan, W. (2018) Time-varying long-term memory in Bitcoin market. *Finance Research Letters*, 25, 280-284. https://doi.org/10.1016/j.frl.2017.12.009

Kim, M. J.; Canh, N. P.; Park, S. Y. (2021) Causal relationship among cryptocurrencies: A conditional quantile approach. *Finance Research Letters*, 42, 101879. https://doi.org/10.1016/j.frl.2020.101879

Koutmos, D. (2018) Return and volatility spillovers among cryptocurrencies. *Economics Letters*, 173, 122-127. https://doi.org/10.1016/j.econlet.2018.10.004

Kristoufek, L. (2018) On Bitcoin markets (in)efficiency and its evolution. *Physica A: Statistical Mechanics and its Applications*, 503, 257-262. https://doi.org/10.1016/j.physa.2018.02.161

Li, Y.; Lucey, B.; Urquhart, A. (2023) Can altcoins act as hedges or safe-havens for Bitcoin? *Finance Research Letters*, 52, 103360. https://doi.org/10.1016/j.frl.2022.103360

Liu, Y.; Li, Z.; Nekhili, R.; Sultan, J. (2023) Forecasting cryptocurrency returns with machine learning. *Research in International Business and Finance*, 64, 101905. https://doi.org/10.1016/j.ribaf.2023.101905

Manahov, V. (2023) The rapid growth of cryptocurrencies: How profitable is trading in digital money? *International Journal of Finance & Economics*. https://doi.org/10.1002/ijfe.2778

Mokni, K.; Ajmi, A. N. (2021) Cryptocurrencies vs. US dollar: Evidence from causality in quantiles analysis. *Economic Analysis and Policy*, 69, 238-252. https://doi.org/10.1016/j.eap.2020.12.011

Nadarajah, S.; Chu, J. (2017) On the inefficiency of Bitcoin. *Economics Letters*, 150, 6-9. https://doi.org/10.1016/j.econlet.2016.10.033

Panagiotidis, T.; Stengos, T.; Vravosinos, O. (2018) On the determinants of Bitcoin returns: A LASSO approach. *Finance Research Letters*, 27, 235-240. https://doi.org/10.1016/j.frl.2018.03.016

Parkinson, M. (1980) The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1), 61-65. https://doi.org/10.1086%2F296071

Politis, D. N.; Romano, J. P. (1994) The stationary bootstrap. *Journal of the American Statistical Association*, 89(428), 1303-1313. https://doi.org/10.1080/01621459.1994.10476870

Politis, D. N.; White, H. (2004) Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, 23(1), 53-70. https://doi.org/10.1081/etc-120028836

Patton, A.; Politis, D. N.; White, H (2009) Correction to "Automatic block-length selection for the dependent bootstrap" by D. Politis and H. White, *Econometric Reviews*, 28(4), 372-375. https://doi.org/10.1080/07474930802459016

Qiao, X.; Zhu, H.; Hau, L. (2020) Time-frequency co-movement of cryptocurrency return and volatility: Evidence from wavelet coherence analysis. *International Review of Financial Analysis*, 71, 101541. https://doi.org/10.1016/j.irfa.2020.101541

Raza, S. A.; Shah, N.; Guesmi, K.; Msolli, B. (2022) How does COVID-19 influence dynamic spillover connectedness between cryptocurrencies? Evidence from non-parametric causality-in-quantiles techniques. *Finance Research Letters*, 47, 102569. https://doi.org/10.1016/j.frl.2021.102569

Sebastião, H. M.; Cunha, P. J.; Godinho, P. M. (2021) Cryptocurrencies and blockchain. Overview and future perspectives. *International Journal of Economics and Business Research*, 21(3), 305-342. https://doi.org/10.1504/IJEBR.2021.114400

Sebastião, H.; Godinho, P. (2021) Forecasting and trading cryptocurrencies with machine learning under changing market conditions. *Financial Innovation*, 7(1), 1-30. https://doi.org/10.1186/s40854-020-00217-x

Shahzad, S. J.; Bouri, E.; Ahmad, T.; Naeem, M. A. (2022) Extreme tail network analysis of cryptocurrencies and trading strategies. *Finance Research Letters*, 44, 102106. https://doi.org/10.1016/j.frl.2021.102106

Shahzad, S. J.; Bouri, E.; Roubaud, D.; Kristoufek, L.; Lucey, B. (2019) Is Bitcoin a better safe-haven investment than gold and commodities? *International Review of Financial Analysis*, 63, 322-330. https://doi.org/10.1016/j.irfa.2019.01.002

Tibshirani, R. (1996) Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Urquhart, A. (2016) The inefficiency of Bitcoin. *Economics Letters*, 148, 80-82. https://doi.org/10.1016/j.econlet.2016.09.019

Wei, W. C. (2018) Liquidity and market efficiency in cryptocurrencies. *Economics Letters*, 168, 21-24. https://doi.org/10.1016/j.econlet.2018.04.00