Article

Contrasting Cryptocurrencies with Other Assets: Full Distributions and the COVID Impact

Esfandiar Maasoumi [1] and Xi Wu [2,]

[1] Department of Economics, Emory University; emaasou@emory.edu
[2] Department of Economics, Emory University; xi.wu@emory.edu

Abstract: We investigate any similarity and dependence based on the full distributions of cryptocurrency assets, stock indices and industry groups. We characterize full distributions with entropies to account for higher moments and non-Gaussianity of returns. Divergence and distance between distributions are measured by metric entropies, and rigorously tested for statistical significance. We assess stationarity and normality of assets, as well as the basic statistics of cryptocurrencies and traditional asset indices, before and after COVID-19 pandemic outbreak. These assessments are not subjected to possible misspecifications of conditional time series models which are also examined for their own interests. We find that NASDAQ daily return has the most similar density and co-dependence with Bitcoin daily return, generally, but after COVID-19 outbreak in early 2020, even S&P500 daily return distribution is statistically closely dependent on, and indifferent from Bitcoin daily return. All asset distances have declined by 75% or more after COVID-19 outbreak. We also find that the highest similarity before COVID-19 outbreak is between Bitcoin and Coal, Steel and Mining industries, and after COVID-19 outbreak is between Bitcoin and Business Supplies, Utilities, Tobacco Products and Restaurants, Hotels, Motels industries, compared to several others. This study shed light on examining distribution similarity and co-dependence between cryptocurrencies and other asset classes, especially demystify effects of the important timely topic, COVID-19.

Keywords: Cryptocurrency, Bitcoin, Entropy, Co-dependence, COVID-19

Citation: Title. *Journal Not Specified* **2021**, *1*, 0. https://doi.org/

Received:
Accepted:
Published:

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Submitted to *Journal Not Specified* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/ 4.0/).

1. Introduction

Since the emergence of Bitcoin based on blockchain technology in 2018, global financial markets have witnessed the birth and rapid rise of cryptocurrencies (cryptos) as a new asset class. Cryptos are based on fundamentally new technologies, the potential of which highly anticipated but not fully understood. In their current form, however, cryptos are also behaving like high growth assets. The cryptocurrency market is an important part of the global assets markets. As of September 2020, there were over 18.53 million Bitcoins in circulation with a total market value of around $199.62 billion.

With the rapid development of cryptocurrency market, the literature has focused on statistical properties and risk behavior of the cryptocurrency in comparison with classical assets, like equities and exchange rates. In the setting of time series models, Pichl and Kaizoji (2017) found that cryptocurrency markets are even more volatile than foreign exchange markets. Chu et al. (2017), Bouri et al. (2017), Katsiampa (2017), Bariviera (2017), Bau et al. (2018) and Stavroyiannis (2018) observed the phenomenon of volatility clustering in cryptocurrency market. Regime-switching behaviors are detected by Bariviera et al. (2017), Balcombe and Fraser (2017). Thies and Molnr (2018) have identified structural breaks in the volatility process of Bitcoin via a Bayesian framework. Lahmiri et al. (2018) and Lahmiri and Bekiros (2018) have pointed out that Bitcoin markets are characterized by long memory and multifractality.

Statistical similarity and co-dependence are central to the analysis of market efficiency and allocation. Most existing studies focus on Bitcoin returns and ”correlation" analysis. For example, Baur et al. (2017) show that Bitcoin returns are essentially uncorrelated with traditional asset classes such as stocks and bonds, which points to diversification possibilities. Other studies investigate the determinants of Bitcoin returns. Li and Wang (2017) suggest that measures of financial and macroeconomic activity are drivers of Bitcoin returns. Kristoufek (2015) considers financial uncertainty, Bitcoin trading volume in Chinese Yuan and Google trends as potential drivers of Bitcoin returns. Recently, many studies examine whether Bitcoin belongs to any existing asset classes, with many comparing it to gold, others to precious metals or to speculative assets (Baur et al. (2017), Bouri et al. (2017)). Some have classified Bitcoin as a new asset class within currency and commodity groups (Dyhrberg (2016)).

Another area of interest is forecasting Bitcoin volatility, since such forecasts represent an important ingredient in risk assessment and allocation, and derivatives pricing theory. Balcilar et al. (2017) analyze the causal relation between trading volume and Bitcoin returns and volatility. They find that volume cannot help to predict the volatility of Bitcoin returns. Bouri et al. (2017) find no evidence for asymmetry in the conditional volatility of Bitcoins when considering the post December 2013 period and investigate the relation between the VIX index and Bitcoin volatility. Al-Khazali et al. (2018) consider a model for daily Bitcoin returns and show that Bitcoin volatility tends to decrease in response to positive news about the US economy.

Scant attention has been paid to the full distributions of these assets. An exception is Osterrieder and Lorenz (2017) and Begusic et al. (2018) who have studied the unconditional distribution of Bitcoin returns and found that it has more probability mass in the tails than that of foreign exchange and stock market returns. Findings that are based on models of return and volatility, possibly with conditional covariates, are in effect assessing if similar mechanisms apply to different asset class returns. While this is an aspect of similarity, it does not respond, and indeed may impinge on the assessment of similarity of return outcomes/ distributions. Similar distributions may arise from different evolutions and mechanisms over time.

Our objective in this paper is to revisit some stylized facts of cryptocurrency markets and employ econometrics models for accurate volatility forecasts. In contrast to previous studies that use time series models to forecast crypto returns, in this paper we use entropy profiles of different asset classes and indices, as well as the cryptos. We test for similarity between cryptocurrency and stock returns in a manner that captures nonlinearities and higher moments, nonparametrically. We consider both Bitcoin and Ethereum, as leading crypto which have large volume and relatively long histories. We use nonparametric entropy metrics to test equality between crypto density and stock market index returns. Time series models (ARIMA and GARCH), in contrast, impose a (traditionally) restrictive linear structure on the return data. This may produce non robust inferences and conclusions.

Efficient market analysis is based on (typically) linear relation between a given asset and market returns. In this paper we examine the general definition of dependence between crypto return and stock market returns. Stochastic independence is tested and degree of dependence is measured with entropy metrics.

The rest of the paper is organized as follows: Section 2 presents the data analysis and some stylized facts. In Section 3, we calculate nonparametric entropy metrics to test the density equality between two cryptos (Bitcoin and Ethereum), two stock market indexes (S&P500 and NASDAQ) and 30 commodity industry groups. We conduct equality tests on both marginal distributions and conditional distributions for two periods (pre-COVID and COVID era) and compare the results. In Section 4, we consider a Diff-in-diff analogy to identify any impact of COVID-19. It is found to be large and significant, producing far greater convergence between asset classes and cryptos. Section 5 provides the concluding remarks.

2. Data and Basic Characteristics

The cryptocurrency data and stock market index data set consists of daily spot exchange rates in units of US dollars are from Yahoo Finance[1]. The price observations of Bitcoin (BTC-USD), Ethereum (ETH-USD), S&P500 stock market index (ˆGSPC) and NASDAQ stock market index (ˆIXIC) range from August 6, 2015 to September 1, 2020. We divided the time period into two parts: pre-COVID (August 6, 2015 – January 31, 2020) and COVID era (February 1, 2020 to September 1, 2020). In each data set of crypto market and stock market index, we have open price, intraday high price, intraday low price, close price (adjusted for splits), adjusted close price (adjusted for both dividends and splits) and volume. To better illustrate the relationship between crypto market data and stock market indexes, we calculate the daily log return using adjusted close price:

$$Return_t = 100 * [ln(P_t) - ln(P_{t-1})], \qquad (1)$$

where P_t denotes the adjusted close price in USD at a time t.

We now document main statistical properties of time series for the returns of S&P500 stock market index, NASDAQ stock market index, Bitcoin and Ethereum. Figure 1 illustrates the time evolution of prices, volumes and daily log-returns for S&P500, NASDAQ, Bitcoin and Ethereum. We notice that both Bitcoin and Ethereum arrive their period specific highest price in December 2017 within our analysis period. After this period price peak, the crypro price dropped dramatically. The descriptive statistics of daily log-returns are reported in Table 1. The daily returns of crypto markets exhibit high variability and excess kurtosis, both during pre-COVID and COVID era periods. The deviations from the Normal distribution are confirmed by the Jarque-Bera test that rejects the null hypothesis of normality.

[1] https://finance.yahoo.com

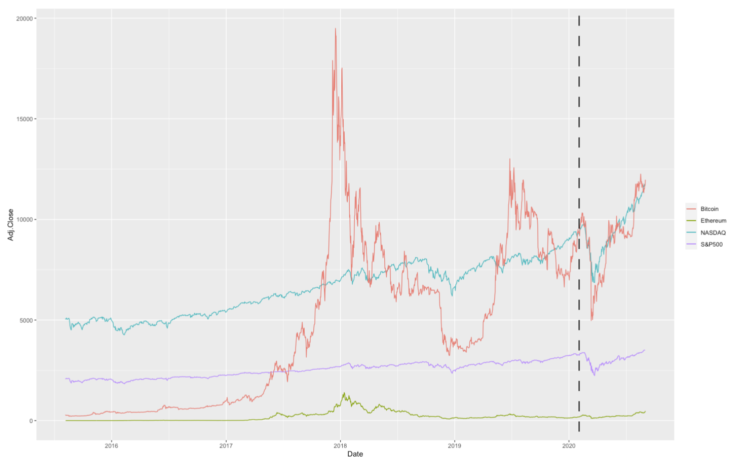
We applied the Augmented-Dicker-Fuller (ADF) unit-root test, which suggests stationarity of the log-returns. An ADF test tests the null hypothesis that a unit root is present in a time series sample. The alternative hypothesis is different depending on which version of the test is used, but is usually stationary or trend-stationary. In our case, we use the alternative hypothesis of stationary. This shows that the null hypothesis is rejected, and the time series of returns in each markets is stationary. These observations suggest that the crypto market is not as efficient as stock or foreign exchange markets, which display a complete lack of predictability (Lahmiri et al. (2018)).

Since early 2020, the COVID-19 wreaked unprecedented havoc on the world economies. Educational institutions, travel industry to public events, almost everything is either postponed or in limbo, which is inevitably going to affect businesses at every turn. Thousands of cases and deaths have already been recorded globally, and with the uncertainty on development of vaccines, the stock markets began to take many hits in terms of new lows. The SP 500 index hit a period low since 2008 when the world plunged into a financial crisis. The cryptocurrency market has even become more volatile and has also experienced one of the worst sudden declines. We also noticed from Figure 1 that both cryptos and stock market indexes became more uncertain since the COVID-19 outbreak in early 2020. The return prices and volumes of Bitcoin and Ethereum also surged since early 2020.

Table 1: Descriptive statistics

	pre-COVID (Aug 2015 - Jan 2020)				COVID era (Feb 2020 - Sep 2020			
Daily log-return	S&P500	Nasdaq	Bitcoin	Ethereum	S&P500	Nasdaq	Bitcoin	Ethereum
Observations	1129	1129	1640	1639	147	147	213	213
Mean	0.04	0.05	0.21	0.25	0.05	0.16	0.11	0.45
Standard deviation	0.86	1.04	3.89	7.09	2.72	2.71	4.61	5.92
Skewness	-0.57	-0.51	-0.18	-3.44	-0.73	-0.92	-4.49	-3.68
Kurtosis	4.12	3.15	4.72	72.46	5.13	5.27	48.02	35.51
Augmented Dickey-Fuller (ADF)	-10.98 **	-11.26 **	-10.93 **	-10.93 **	-5.64 **	-5.48 **	-5.16 **	-4.98 **
Jarque-Bera	862.50 ***	518.27 ***	1538.80 ***	362486 ***	180.51 ***	197.22 ***	21507 ***	11855 ***

*Note: Entries marked with *** have empirical p-values* < 0.01, ** $0.01 \leq p < 0.05$, *and* * $0.05 \leq p < 0.10$ *under the null of non-stationary data for ADF test and the null of normally distributed data for Jarque-Bera test.*



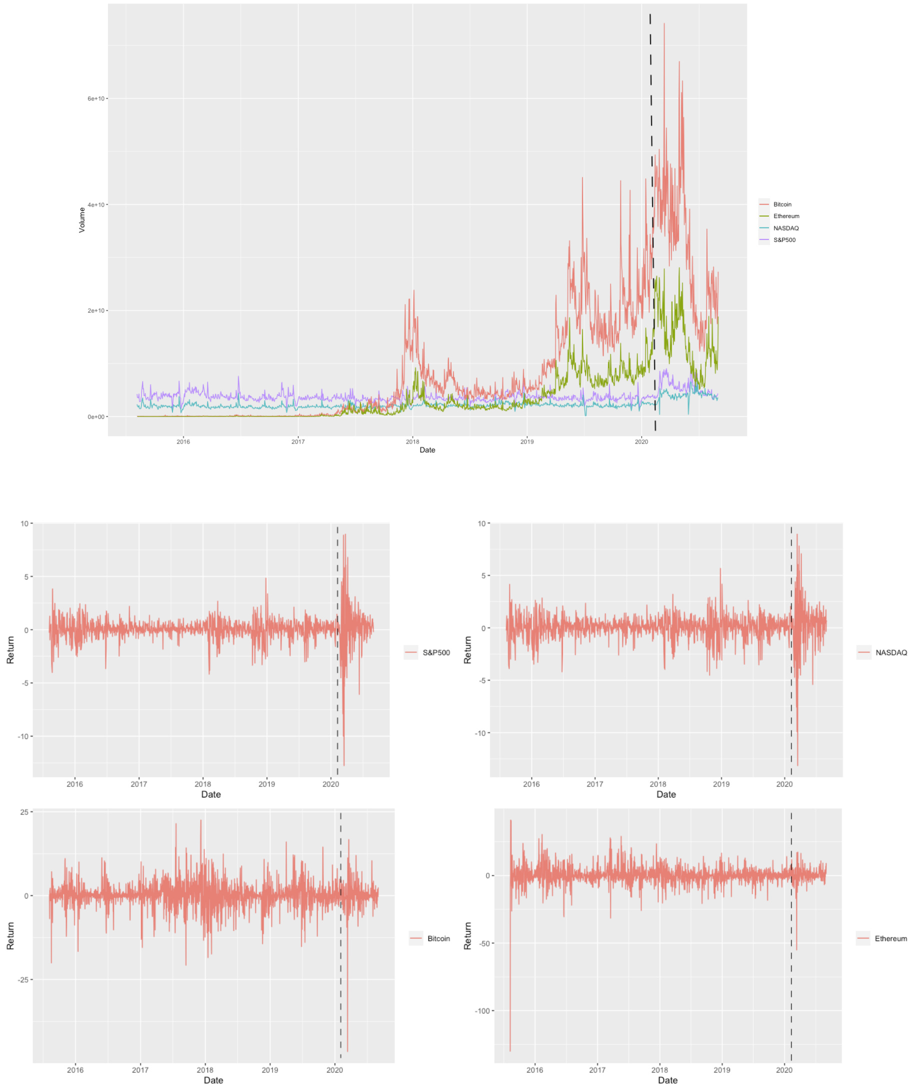


Figure 1. Plot of price, volume and daily log-returns

3. Entropy Profiles Method

3.1. Brief Introduction to Information Theory and Entropy

Consider two variables X and Y. Correlation between them may be ill defined when they are discrete, and may be a poor measure of "relation" when nonlinearity and/or non-Gaussianity is involved.

Let $\Re = \{a_1, a_2, ..., a_M\}$ be a finite set and p be a proper probability mass function (PDF) on $\Re$. The amount of information needed to fully characterize all of the elements of this set consisting of M discrete elements is defined by $I(\Re_M) = log_2 M$ and is known as Hartley's formula. Shannon (1948) built on Hartley's formula in the context of digitization and communications, to develop Shannon's entropy:

$$H(p) = -\sum_{i=1}^{M} p_i log(p_i), \tag{2}$$

with $xlog(x)$ tending to zero as x tends to zero. This information criterion measures the uncertainty or informational content that is implied by p. The entropy-uncertainty measure $H(p)$ reaches a maximum when $p_1 = p_2 = ... = p_M = 1/M$ (and is equal to Hartley's formula) and a minimum with a point mass function. It is emphasized here that $H(p)$ is a function of the probability distribution. For example, if η is a random variable with possible distinct realizations $x_1, x_2, ..., x_M$ with probabilities $p_1, p_2, ..., p_M$, the entropy $H(p)$ does not depend on the values $x_1, x_2, ..., x_M$ of η. If, on the other hand, η is a continuous random variable, then the entropy of a continuous density is

$$H(x) = -\int p(x)log(p(x))dx, \tag{3}$$

a differential entropy.

Renyi (1961, 1970) showed that, for a (sufficiently often) repeated experiment, one needs on average the amount $H(p) + \epsilon$ of zero-one symbols (for any positive ϵ) in order to characterize an outcome of that experiment. Thus, it seems logical to "expect" that the outcome of an experiment contains $H(p)$ information.

Similarly, $H(p)$ is a measure of uncertainty about a specific possible outcome before observing it, which is equivalent to the amount of randomness represented by p. It is proportional to "variance" in the case of a Normal distribution. Thus entropy is a far superior and robust measure of volatility/risk than variance for non Gaussian phenomena. It is indeed unique for any distribution, much as the characteristic function is, both representing all the moments of a distribution, which could be merely the mean and variance in the case of a Normal variable. Asset returns are not Gaussian!

Given a prior or competing distribution q, defined on $\Re$, the cross-entropy (Kullback-Leibler, K-L, 1951) measure is

$$I(p;q) = \sum_{i=1}^{M} p_i log(p_i/q_i), \tag{4}$$

where a uniform q reduces $I(p;q)$ to $H(p)$. This measure reflects the gain in information with respect to $\Re$ resulting from the additional knowledge in p relative to q. Like with $H(p)$, $I(p;q)$ is an information theoretic distance of p from q. It can be symmetrized by averaging $I(p;q)$ and $I(q;p)$.

Facing the fundamental question of drawing inferences from limited and insufficient data, Jaynes proposed the maximum entropy (ME) principle, which he viewed as a generalization of Bernoulli and Laplace's Principle of Insufficient Reason.

Given T constraints, perhaps in the form of moments, Jaynes proposed the ME method, which is to maximize $H(p)$ subject to the T structural constraints. Thus, given moment conditions, X_t $(t = 1, 2, ..., T)$, where $T < M$, the ME principle prescribes choosing the $p(a_i)$ that maximizes $H(p)$ subject to the given constraints (moments) of the problem. The solution to this underdetermined problem is

$$\widehat{p(a_i)} \propto exp\{-\sum_t \hat{\lambda}_t X_t(a_i)\}, \tag{5}$$

where λ are the T Lagrange multipliers, and $\hat{\lambda}$ are the values of the optimal solution (estimated values) of λ. Naturally, if no constraints are imposed, $H(p)$ reaches its maximum value and the p are distributed uniformly.

Building on Shannon's work, a number of generalized entropies and information measures were developed. Starting with the idea of describing the gain of information, Renyi (1970) developed the entropy of order α for incomplete random variables. The relevant generalized entropy measure of a proper probability distribution is

$$H_{\alpha}^{R}(p) = \frac{1}{1-\alpha} log \sum_{k} p_k^{\alpha}. \tag{6}$$

Shannon measure is a special case of this measure where $\alpha \to 1$. Similarly, the Renyi cross-entropy of order α is

$$I_{\alpha}^{R}(x|y) = I_{\alpha}^{R}(p,q) = \frac{1}{1-\alpha} log \sum_{k} \frac{p_k^{\alpha}}{q_k^{\alpha-1}}, \tag{7}$$

which is equal to the traditional cross-entropy measure as $\alpha \to 1$. Only one member of these ″divergence" measures is a metric, which we define below.

Entropy has been actively considered in finance theory since at least 1999. According to Gulko (1999), ″entropy pricing theory" suggests that in information efficient markets, perfectly uncertain market beliefs must prevail. Using entropy to measure market uncertainty, entropy-maximizing market beliefs must prevail. One can derive (entropy) optimal asset pricing models that are similar to Black-Scholes model (with the log-normal distribution replaced by other probability distributions).

3.2. Using entropy to test equality of univariate densities

Maasoumi & Racine (2002) considered a metric entropy that is useful for testing for equality of densities for two univariate random variables X and Y. The function computes the nonparametric metric entropy (normalized Hellinger, or Granger et al. (2004)) for testing the null of equality of two univariate density (or probability) functions. For continuous variables,

$$\begin{aligned} S_{\rho} &= \frac{1}{2} \int (f_1^{1/2} - f_2^{1/2})^2 dx \\ &= \frac{1}{2} \int (1 - \frac{f_2^{1/2}}{f_1^{1/2}})^2 dF_1(x), \end{aligned} \tag{8}$$

where $f_1 = f(x)$ and $f_2 = f(y)$ are the marginal densities of the random variables X and Y. The second expression is in a moment from which is often replaced with a sample average, especially for theoritical developments. If the density of X and the density of Y are equal, this metric will yield the value zero, and is otherwise positive and less than one. We use S_{ρ} to test the distance between crypto density and stock market index density. Some properties this entropy measure S_{ρ} are given in (Granger et al. (2000)), and Gianerinni, Maasoumi and Dagum (2015). In particular, the modulus of S_{ρ} is between 0 and unity; S_{ρ} is equal to or has a simple relationship with the (linear) correlation coefficient in the case of a bivariate normal distribution; S_{ρ} is metric, that is, it is a true measure of distance and not just of ″divergence". This is especially important in our applications where triangularity property is required in meaningful comparative assessments of several distances and asset classes.

Software for nonparametric kernel smoothing implementation of this metric is made available in R (NP package) among others. For the kernel function, we employ the widely used nonparametric second-order Gaussian kernel, while bandwidths are selected via likelihood cross-validation (Silverman (1986)). Bootstrap is conducted via resampling with replacement from the pooled empirical distributions of X and Y under the null hypothesis of equality.

We estimate the metric S_ρ for the daily returns data for $x = Return_{crypto}$ and $y = Return_{stock}$. Table 2 shows the S_ρ values and the corresponding p-values. As was noted in Granger et al. (2000) and Skaug & Tjostheim (1996), the asymptotic distribution of S_ρ is unreliable for practical inference, We therefore compute p-values by resampling the statistic under the null of equality.

Examining Table 2 we see that S_ρ is smallest between x = Bitcoin and y = NASDAQ, both during pre-COVID and COVID era periods, which indicates that the distance between the densities of Bitcoin daily returns and NASDAQ daily returns is smaller than other combinations. The p-value shows that the result is significant. By visualizing the result in Figure 2 - Figure 5, we can also see the Bitcoin daily returns density and the NASDAQ stock market index daily returns density have similar shapes. While during COVID era, also S&P500 returns distribution is statistically closely dependent on, and indifferent from Bitcoin's.

Comparing S_ρ before and after the COVID-19 outbreak, we conclude that the values of S_ρ decrease generally in all cases, sometimes dramatically. This suggests that the densities of crypto and stock index returns became more similar with the advent of COVID-19. This mostly due to a large change in the distribution of major stock indices, but also party due to a smaller movement in crypto distributions.

Table 3 reveals the entropy metric S_ρ of the assets themselves pre-COVID & COVID era. By doing so, we can see if the difference between the cryptos and stocks is partly due to specific asset change caused by the effect of COVID-19. The results show that the distributions of S&P500 and NASDAQ changed dramatically and significantly before and after COVID-19 outbreak, which indicates that the changes of S_ρ between cryptos and stocks may mainly caused by the changes of stocks' distributions. We will dive deeper on this part in Section 4.

Table 2: Test equality of univariate densities: cryptos & stocks

	pre-COVID (Aug 2015 - Jan 2020)		COVID era (Feb 2020 - Sep 2020)		Difference
Daily log-return	S_rho	p-value	S_rho	p-value	
S&P500 & Bitcoin	0.20	2.22e-16 ***	0.04	0.1010	-0.16
S&P500 & Ethereum	0.33	2.22e-16 ***	0.08	2.22e-16 ***	-0.25
NASDAQ & Bitcoin	0.16	2.22e-16 ***	0.04	0.0404 *	-0.12
NASDAQ & Ethereum	0.28	2.22e-16 ***	0.08	2.22e-16 ***	-0.20

*Note: Entries marked with *** have empirical p-values* < 0.01, ** $0.01 \leq p < 0.05$, *and* * $0.05 \leq p < 0.10$ *under the null of independence of returns.*

Table 3: Test equality of univariate densities: assets with themselves pre-COVID & COVID era

Daily log-return	S_rho	p-value
S&P500 with itself pre-COVID & COVID era	0.13	<2.22e-16 ***
NASDAQ with itself pre-COVID & COVID era	0.10	<2.22e-16 ***
Bitcoin with itself pre-COVID & COVID era	0.02	0.3737
Ethereum with itself pre-COVID & COVID era	0.02	0.0303 *

*Note: Entries marked with *** have empirical p-values* < 0.01, ** $0.01 \le p < 0.05$, *and* * $0.05 \le p < 0.10$ *under the null of independence of returns.*

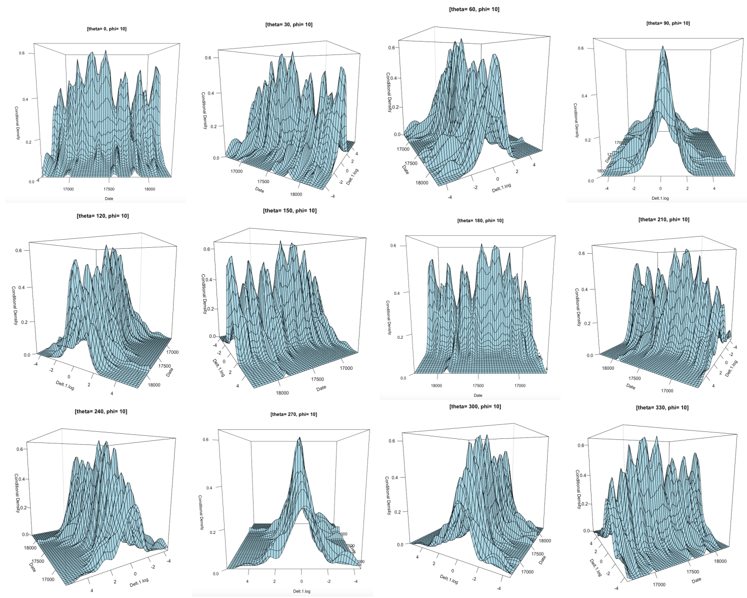


Figure 2. Density of NASDAQ: pre-COVID

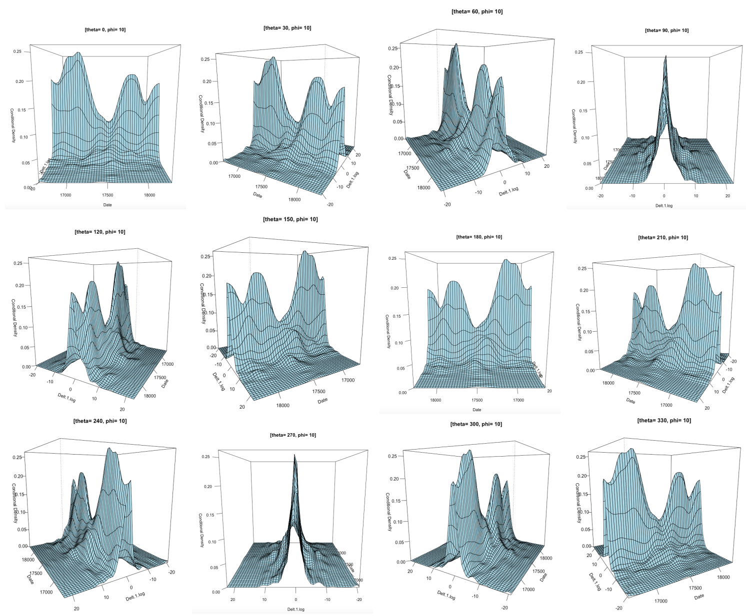


Figure 3. Density of Bitcoin: pre-COVID

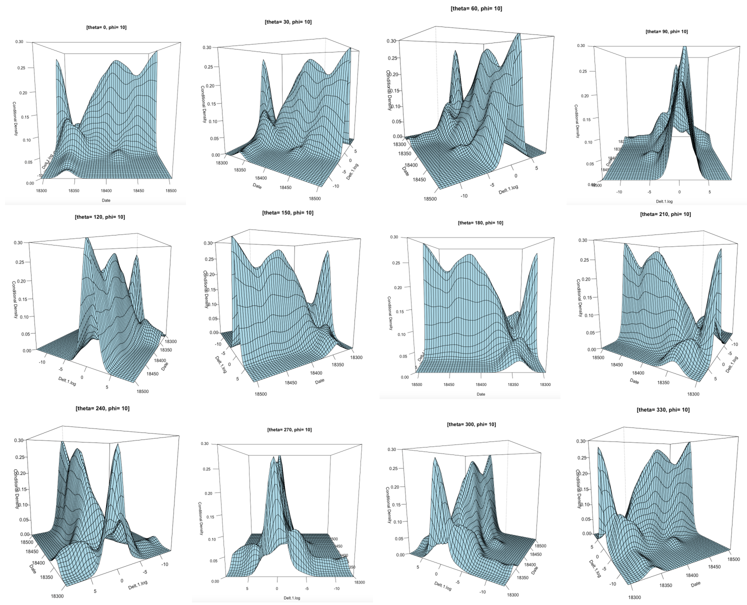


Figure 4. Density of NASDAQ: COVID era

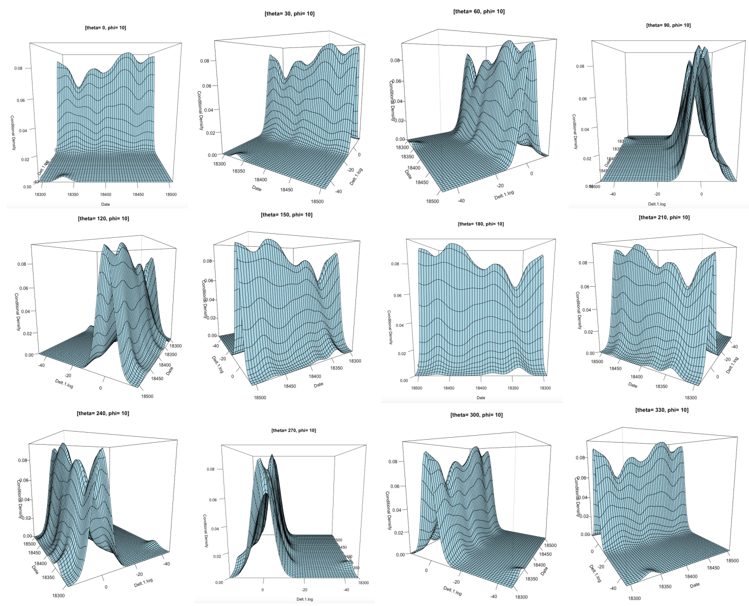


Figure 5. Density of Bitcoin: COVID era

3.3. Similarity with Select Asset Classes

In this part, we apply the same method to test the equality of densities for daily returns of Bitcoin and stocks in different industry groups. The data for daily stock returns in different industries comes from Kenneth R. French 30 Industry Portfolios [2]. The Kenneth R. French 30 Industry Portfolios data set was created by *CMPT_IND_RETS_DAILY* using the 202006 CRSP database, assigned each NYSE, AMEX, and NASDAQ stock to an industry portfolio at the end of June of year t based on its four-digit SIC code at that time, then computed returns from July of t to June of

[2] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data$_L ibrary/det_3 0_i nd_p ort.html$

$t + 1$. We use the daily average value weighted returns for 30 industry portfolios data. The 30 industry portfolios include: Food Products (Food), Beer Liquor (Beer), Tobacco Products (Smoke), Recreation (Games), Printing and Publishing (Books), Consumer Goods (Hshld), Apparel (Clths), Healthcare (Hlth), Medical Equipment, Pharmaceutical Products, Chemicals (Chems), Textiles (Txtls), Construction and Construction Materials (Cnstr), Steel Works Etc (Steel), Fabricated Products and Machinery (Fabpr), Electrical Equipment (Elceq), Automobiles and Trucks (Autos), Aircraft, ships, and railroad equipment (Carry), Precious Metals, Non-Metallic, and Industrial Metal Mining (Mines), Coal (Coal), Petroleum and Natural Gas (Oil), Utilities (Util), Communication (Telcm), Personal and Business Services (Servs), Business Equipment (Buseq), Business Supplies and Shipping Containers (Paper), Transportation (Trans), Wholesale (Whlsl), Retail (Rtail), Restaraunts, Hotels, Motels (Meals), Banking, Insurance, Real Estate, Trading (Fin), Everything Else (Other). We apply the nonparametric entropy metrics test of equality of densities proposed in Maasoumi & Racine (2002), described above, where $f_1 = f(x)$ and $f_2 = f(y)$ are the marginal densities of daily returns of Bitcoin and stocks in different industries, respectively.

From Table 4, we calculated the entropy measures between Bitcoin and select asset classes. During pre-COVID period, the density of Bitcoin daily return has smallest distance with the density of Coal industry daily return. The S_ρ between these two densities is 0.02 and statistically significant. The density of Bitcoin daily return also has small distances with densities of Steel Works Etc, as well as Precious Metals, Non-Metallic, and Industrial Metal Mining industries, with S_ρ values of 0.07 and 0.09 respectively. During COVID era, the density of Bitcoin daily return has smallest distance with the density of Business Supplies and Shipping Containers, Utilities, Tobacco Products and Restaraunts, Hotels, Motels industries daily returns, with S_ρ values of 0.03. Comparing S_ρ before and after the COVID-19 outbreak, we conclude that the values of S_ρ decrease generally in all cases. This is consistent with our findings with stock indexes in the previous section, which indicates that forecasting cryptos' performance could be more feasible during COVID era.

We also calculated the S_ρ with select asset classes with themselves before and after the COVID-19 outbreak (see column 2 in Table 4). It is clear that for all industry groups during COVID era, the asset distributions diverge from their own pre-COVID distributions, and the distribution divergence of industry groups are more significant comparing with cryptos' (shown in Table 3).

Table 4: Entropy measure between Bitcoin and different Industries

	pre-COVID and COVID era with itself		pre-COVID with Bitcoin		COVID era with Bitcoin		Difference
Daily log-return	S_rho	p-value	S_rho	p-value	S_rho	p-value	
Food	0.16	<2.22e-16 ***	0.22	<2.22e-16 ***	0.04	0.0808 .	-0.18
Beer	0.14	<2.22e-16 ***	0.21	<2.22e-16 ***	0.07	0.1010	-0.14
Smoke	0.14	<2.22e-16 ***	0.14	<2.22e-16 ***	0.03	0.2121	-0.11
Games	0.09	<2.22e-16 ***	0.10	<2.22e-16 ***	0.05	0.0202 *	-0.05
Books	0.19	<2.22e-16 ***	0.15	<2.22e-16 ***	0.04	0.0909 .	-0.11
Hshld	0.14	<2.22e-16 ***	0.21	<2.22e-16 ***	0.04	0.4040	-0.17
Clths	0.20	<2.22e-16 ***	0.12	<2.22e-16 ***	0.04	0.1212	-0.08
Hlth	0.12	<2.22e-16 ***	0.17	<2.22e-16 ***	0.04	0.1717	-0.13
Chems	0.21	<2.22e-16 ***	0.15	<2.22e-16 ***	0.04	0.1414	-0.11
Txtls	0.26	<2.22e-16 ***	0.11	<2.22e-16 ***	0.07	0.0101 *	-0.04
Cnstr	0.23	<2.22e-16 ***	0.14	<2.22e-16 ***	0.04	0.2020	-0.10
Steel	0.14	<2.22e-16 ***	0.07	<2.22e-16 ***	0.05	0.0202 *	-0.02
Fabpr	0.19	<2.22e-16 ***	0.13	<2.22e-16 ***	0.04	0.0808 .	-0.09
Elceq	0.22	<2.22e-16 ***	0.14	<2.22e-16 ***	0.04	0.1111	-0.10
Autos	0.21	<2.22e-16 ***	0.12	<2.22e-16 ***	0.04	0.1212	-0.08
Carry	0.27	<2.22e-16 ***	0.15	<2.22e-16 ***	0.06	0.0202 *	-0.08
Mines	0.09	<2.22e-16 ***	0.09	<2.22e-16 ***	0.05	0.0505 .	-0.05
Coal	0.09	<2.22e-16 ***	0.02	<2.22e-16 ***	0.09	<2.22e-16 ***	0.07
Oil	0.22	<2.22e-16 ***	0.11	<2.22e-16 ***	0.05	0.0101 *	-0.05
Util	0.22	<2.22e-16 ***	0.22	<2.22e-16 ***	0.03	0.3939	-0.18
Telcm	0.19	<2.22e-16 ***	0.20	<2.22e-16 ***	0.04	0.1313	-0.16
Servs	0.14	<2.22e-16 ***	0.16	<2.22e-16 ***	0.05	0.1111	-0.11
Buseq	0.13	<2.22e-16 ***	0.14	<2.22e-16 ***	0.04	0.1717	-0.10
Paper	0.17	<2.22e-16 ***	0.18	<2.22e-16 ***	0.03	0.3535	-0.15
Trans	0.18	<2.22e-16 ***	0.15	<2.22e-16 ***	0.04	0.1515	-0.11
Whlsl	0.24	<2.22e-16 ***	0.19	<2.22e-16 ***	0.04	0.2020	-0.15
Rtail	0.10	<2.22e-16 ***	0.18	<2.22e-16 ***	0.08	<2.22e-16 ***	-0.10
Meals	0.24	<2.22e-16 ***	0.20	<2.22e-16 ***	0.03	0.2626	-0.17
Fin	0.25	<2.22e-16 ***	0.16	<2.22e-16 ***	0.05	0.1010	-0.11
Other	0.20	<2.22e-16 ***	0.20	<2.22e-16 ***	0.04	0.1010	-0.16

*Note: Entries marked with *** have empirical p-values* < 0.01, ** $0.01 \le p < 0.05$, *and* * $0.05 \le p < 0.10$ *under the null of independence of returns.*

3.4. Testing General Nonlinear Co-dependence

The above test of Maasoumi and Racine (2002) may be employed for testing stochastic independence of any two random variables X and Y. Let $f_1 = f(x_i, y_i)$ be the joint density and $f_2 = g(x_i) * h(y_i)$ be the product of the marginal densities. The unknown density functions are replaced with nonparametric kernel estimates. The methodology is as before, with the null of independence imposed in the bootstrap resampling implementation of the test. Bandwidths are obtained via likelihood cross-validation by default for the marginal and joint densities.

The results are in Table 5. There is significant dependence only between Bitcoin and NASDAQ before COVID-19 outbreak. During COVID era, independence is comfortably rejected for all pairings. The two situations represent very radical changes in the status of cryptos for portfolio diversification.

Table 5: Independence test

	pre-COVID (Aug 2015 - Jan 2020)		COVID era (Feb 2020 - Sep 2020)		Difference
Daily log-return	S_rho	p-value	S_rho	p-value	
S&P500 & Bitcoin	0.0085	0.0303 *	0.0148	2.22e-16 ***	0.0063
S&P500 & Ethereum	0.0076	0.5758	0.0172	2.22e-16 ***	0.0096
NASDAQ & Bitcoin	0.0072	0.0101 *	0.0163	2.22e-16 ***	0.0091
NASDAQ & Ethereum	0.0061	0.6061	0.0178	2.22e-16 ***	0.0117

*Note: Entries marked with *** have empirical p-values* < 0.01, ** $0.01 \le p < 0.05$, *and* * $0.05 \le p < 0.10$ *under the null of independence of returns.*

4. Difference-in-differences analysis

Difference in differences (Diff-in-diff) is a statistical technique used in econometrics and quantitative research that attempts to mimic an experimental research design using observational study data, by studying the differential effect of a treatment on a ”treatment group" versus a ”control group" in a natural experiment. It calculates the effect of a treatment on an outcome by comparing the average change over time in the outcome variable for the treatment group, compared to the average change over time for the control group.

Before we construct our Diff-in-diff model, we would like to emphasize that the entropy metrics exhibit linear decomposition property. The reason why we can decompose S_ρ is that it is a metric, which means it satisfies the triangularity property of distances. Therefore, we can write the entropy metric between stock and crypto during COVID era as the summation of the entropy metric between them during pre-COVID period plus a time trend λ_t and plus the COVID effect.

$$S_\rho(f_{s_i,t_2}, f_{c_j,t_2}) = S_\rho(f_{s_i,t_1}, f_{c_j,t_1}) + \lambda_t + COVID + \epsilon_{i,j}, \tag{9}$$

where $S_\rho(f_{s_i,t_2}, f_{c_j,t_2})$ stands for the entropy metric between stock i and crypto j during COVID era, and $S_\rho(f_{s_i,t_1}, f_{c_j,t_1})$ stands for the entropy metric between stock i and crypto j during pre-COVID period. λ_t is the time trend defined by $\lambda_t = S_\rho(f_{s_i,t_2}, f_{s_i,t_1}) + S_\rho(f_{c_j,t_2}, f_{c_j,t_1})$, which measures the entropy metric of both stock i and crypto j from pre-COVID period to COVID era with itself. $COVID$ is the effect of exogenous shock provided by COVID-19 to the entropy metrics. $\epsilon_{i,j}$ is the residual term.

Since we have already calculated the distribution distances between assets in the previous sections, from equation (9), we can easily estimate the COVID effect on the entropy metrics, say $\widehat{COVID}$. Using entropy metrics S_ρ between Bitcoin and other assets (including S&P500, NASDAQ, the the 30 industry portfolios), we can estimate the COVID effect $\widehat{COVID} = -0.30$. This indicates that after the broke out of COVID-19 pandemic, the distributions of stocks and cryptos became more similar and less independent, quantitatively, the entropy metrics decrease by -0.30 in average.

Next, we follow Card Krueger (1994) to construct our Diff-in-diff model:

$$S_\rho(f_{A_i,t_j}, f_0) = \beta_0 + \beta_1 * Covid + \beta_2 * Crypto + \beta^{DID} * (Covid * Crypto) + \epsilon, \tag{10}$$

where the dependent variable $S_\rho(f_{A_i,t_j}, f_0)$ is our variable of interest, it stands for the entropy metric between asset i's distribution at time j, f_{A_i,t_j}, and a benchmark distribution f_0. $Crypto$ and $Covid$ are dummy variables. $Crypto$ equals to 1 if the asset is crypto, while it equals to 0 if the asset is stock. $Covid$ equals to 1 if during the COVID era and it equals to 0 if during the pre-COVID period. The coefficient for the interaction term, $Covid * Crypto$, is the Diff-in-diff estimator. In this way, we construct our Diff-in-diff model for entropy metric.

We come up with a new method to use our nonparametric entropy metric to estimate the Diff-in-diff estimator. In Table 6, we show the decomposition of the Diff-in-diff analysis. The reason why we can decompose S_ρ is that it is a metric, which means it satisfies the triangularity property of distances. If you take three points, A, B and C, the distance between any of those points is smaller than the total of the other two distances. Also note that S_ρ is a ”squared integral". The second line in Equation (8) also tells us that it is a simple expectation of $1 - (f_2/f_1)^{1/2}$. This is equal to metric developed by Bhathacharya as a measure of relations between two variables. By algebra, we can derive the Diff-in-diff estimator as the entropy metrics between stocks and crytos during COVID era subtract the entropy metric between them during pre-COVID period: $\hat{\beta}^{DID} = S_\rho(f_{s_i,t_2}, f_{c_j,t_2}) - S_\rho(f_{s_i,t_1}, f_{c_j,t_1})$.

Table 6: DID decomposition

Distribution	Stock	Crypto	Difference
pre-COVID	$S_\rho(f_{s_i,t_1}, f_0)$	$S_\rho(f_{c_j,t_1}, f_0)$	$S_\rho(f_{s_i,t_1}, f_{c_j,t_1})$
COVID era	$S_\rho(f_{s_i,t_2}, f_0)$	$S_\rho(f_{c_j,t_2}, f_0)$	$S_\rho(f_{s_i,t_2}, f_{c_j,t_2})$
Change	$S_\rho(f_{s_i,t_2}, f_{s_i,t_1})$	$S_\rho(f_{c_j,t_2}, f_{c_j,t_1})$	$S_\rho(f_{s_i,t_2}, f_{c_j,t_2}) - S_\rho(f_{s_i,t_1}, f_{c_j,t_1})$

5. Conclusion

This paper investigates the similarity and co-dependence between cryptocurrencies daily returns and stock daily returns, before and after the COVID-19 outbreak in early 2020.

Data exhibited different features before and after COVID-19 outbreak. There is similarity between Bitcoin and NASDAQ stock market index with or without the COVID event. The similarity and dependence between cryptos and stock market indexes has become stronger after COVID-19 outbreak. Our findings are robust to model misspecification, and avoid linear measures of dependence and correlation. The entropy profiles method and time series models play different roles in forecasting cryptocurrency returns volatility, and these approaches are complimentary. The time series models elaborate the dynamic movement of returns, on average (conditional mean models). The entropy profiles method is a nonparametric approach which reveals the evolution of the entire distributions and their quantiles. In this paper, we have several findings: Firstly, we found that during pre-COVID period, NASDAQ return and Bitcoin return's distributions are the most similar. Secondly, we can see during the COVID era, the distances between all asset returns have declined by 75% or more, and most of these changes are caused by changes of stock return distributions. We also found that the asset group with the closest similarity with Bitcoin are Coal, Steel and Mining industries during pre-COVID period, and Business Supplies, Utilities, Tobacco Products and Restaurants, Hotels, Motels industries, compared to several others during COVID era. Finally, through non-linear co-dependence test, we found that during COVID era, the densities of stocks and cryptos became more similar and less independent. These results are meaningful because we revealed the similarity and dependence structure between crypto and stock distributions. This can be useful in applying existing theories on stocks to cryptos.

As for future directions of this study, we plan to examine newer data as we have observe the effective vaccines rollout, stock market volatility and the crypto prices peak to new high in 2021. We believe the examination of newer data will drive more promising and effective policy implications.

References

1. Ante, L., Sandner, P., Fiedler, I. (2018). Blockchain-based ICOs: Pure hype or the dawn of a new era of startup financing?. *Journal of Risk and Financial Management*, 11(4), 80.
2. Backus, D., Chernov, M., Zin, S. (2014). Sources of entropy in representative agent models. *The Journal of Finance*, 69(1), 51-99.

3. Baker, S. R., Bloom, N., Davis, S. J., Kost, K. J., Sammon, M. C., Viratyosin, T. (2020). The unprecedented stock market impact of COVID-19 (No. w26945). National Bureau of Economic Research.
4. Baldwin, R., Di Mauro, B. W. (2020). Economics in the time of COVID-19: A new eBook. *VOX CEPR Policy Portal*, 2-3.
5. Bera, A. K., Park, S. Y. (2008). Optimal portfolio diversification using the maximum entropy principle. *Econometric Reviews*, 27(4-6), 484-512.
6. Bocart, F. (2018). Inflation propensity of Collatz orbits: a new proof-of-work for blockchain applications. *Journal of Risk and Financial Management*, 11(4), 83.
7. Borodin, A., El-Yaniv, R., Gogan, V. (2004). Can we learn to beat the best stock. *Journal of Artificial Intelligence Research*, 21, 579-594.
8. Chen, C. Y. H., Hafner, C. M. (2019). Sentiment-induced bubbles in the cryptocurrency market. *Journal of Risk and Financial Management*, 12(2), 53.
9. Conrad, C., Custovic, A., Ghysels, E. (2018). Long-and short-term cryptocurrency volatility components: A GARCH-MIDAS analysis. *Journal of Risk and Financial Management*, 11(2), 23.
10. Fischer, T. G., Krauss, C., Deinert, A. (2019). Statistical arbitrage in cryptocurrency markets. *Journal of Risk and Financial Management*, 12(1), 31.
11. Giannerini, S., Maasoumi, E., Dagum, E. B. (2007). Entropy testing for nonlinearity in time series. Bulletin of the International Statistical Institute, 56th session. ISI.
12. Giannerini, S., Maasoumi, E., Dagum, E. B. (2015). Entropy testing for nonlinear serial dependence in time series. *Biometrika*, 102(3), 661-675.
13. Golan, A. (2002). Information and entropy econometrics—editor's view. *Journal of econometrics*, 107(1-2), 1-15.
14. Gu, R., Shao, Y. (2016). How long the singular value decomposed entropy predicts the stock market?—Evidence from the Dow Jones Industrial Average Index. *Physica A: Statistical Mechanics and its Applications*, 453, 150-161.
15. Henriques, I., Sadorsky, P. (2018). Can bitcoin replace gold in an investment portfolio?. *Journal of Risk and Financial Management*, 11(3), 48.
16. Jana, P., Roy, T. K., Mazumder, S. K. (2007). Multi-objective mean-variance-skewness model for portfolio optimization. *Advanced Modeling and Optimization*, 9(1), 181-193.
17. Jotanovic, V., D'Ecclesia, R. L. (2019). Do Diamond Stocks Shine Brighter than Diamonds?. *Journal of Risk and Financial Management*, 12(2), 79.
18. Kjærland, F., Khazal, A., Krogstad, E. A., Nordstrøm, F. B., Oust, A. (2018). An analysis of bitcoin's price dynamics. *Journal of Risk and Financial Management*, 11(4), 63.
19. Li, Q., Maasoumi, E., Racine, J. S. (2009). A nonparametric test for equality of distributions with mixed categorical and continuous data. *Journal of Econometrics*, 148(2), 186-200.
20. Liu, Y., Tsyvinski, A. (2021). Risks and returns of cryptocurrency. *The Review of Financial Studies*, 34(6), 2689-2727.
21. Maasoumi, E., Eren, O. (2006). The information basis of matching with propensity score (No. 0606).
22. Maasoumi, E., Racine, J. (2002). Entropy and predictability of stock market returns. *Journal of Econometrics*, 107(1-2), 291-312.
23. Maasoumi, E., Zandvakili, S. (1990). Generalized entropy measures of mobility for different sexes and income levels. *Journal of Econometrics*, 43(1-2), 121-133.
24. Maloney, W. F., Taskin, T. (2020). Determinants of social distancing and economic activity during COVID-19: A global view. World Bank Policy Research Working Paper, (9242).
25. O'Dwyer, K. J., Malone, D. (2014). Bitcoin mining and its energy footprint.
26. Pele, D. T., Mazurencu-Marinescu-Pele, M. (2019). Using high-frequency entropy to forecast Bitcoin's daily Value at Risk. *Entropy*, 21(2), 102.

27. Ramelli, S., Wagner, A. F. (2020). Feverish stock price reactions to COVID-19. *The Review of Corporate Finance Studies*, 9(3), 622-655.
28. Sandoval, L. (2014). Structure of a global network of financial companies based on transfer entropy. *Entropy*, 16(8), 4443-4482.
29. Schnaubelt, M., Rende, J., Krauss, C. (2019). Testing stylized facts of bitcoin limit order books. *Journal of Risk and Financial Management*, 12(1), 25.
30. Shintate, T., Pichl, L. (2019). Trend prediction classification for high frequency bitcoin time series with deep learning. *Journal of risk and Financial management*, 12(1), 17.
31. Soest, A. V., Das, M., Gong, X. (2002). A structural labour supply model with flexible preferences. *Journal of Econometrics*, 107(1–2), 345-374.
32. Velde, F. R. (2020). What happened to the US economy during the 1918 influenza pandemic? A view through high-frequency data.
33. von Furstenberg, G. M., Maasoumi, E. (1988). Macroeconomic Implications of the Information Revolution. *The American Economic Review*, 78(2), 178-181.
34. Yu, J. R., Lee, W. Y., Chiou, W. J. P. (2014). Diversified portfolios with different entropy measures. *Applied Mathematics and Computation*, 241, 47-63.
35. Zunino, L., Zanin, M., Tabak, B. M., Pérez, D. G., Rosso, O. A. (2009). Forbidden patterns, permutation entropy and stock market inefficiency. *Physica A: Statistical Mechanics and its Applications*, 388(14), 2854-2864.
36. Zunino, L., Zanin, M., Tabak, B. M., Pérez, D. G., Rosso, O. A. (2010). Complexity-entropy causality plane: A useful approach to quantify the stock market inefficiency. *Physica A: Statistical Mechanics and its Applications*, 389(9), 1891-1901.