Distribution Agreement

Essays on Diversification: Cryptocurrencies and Other Assets

By

Xi Wu
Doctor of Philosophy

Economics

Esfandiar Maasoumi, Ph.D.
Advisor

Zhongjian Lin, Ph.D.
Committee Member

Ruixuan Liu, Ph.D.
Committee Member

Accepted:

Kimberly Jacob Arriola, Ph.D, MPH
Dean of the James T. Laney School of Graduate Studies

Date

Essays on Diversification: Cryptocurrencies and Other Assets

By

Xi Wu
M.S., Massachusetts Institute of Technology, 2015
B.S., Nankai University, 2014

Advisor: Esfandiar Maasoumi, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2022

Abstract

Essays on Diversification: Cryptocurrencies and Other Assets
By Xi Wu

There are three essays in this doctoral dissertation that study assets similarity and investment diversification. The purpose of this dissertation is to make clear a central theme: Diversification makes it desirable to identify new asset classes. The decision of what is a new asset class is based on its comparison with other assets. This comparison can be based on traditional methods which compare predictive models, such as GARCH and other time series models, summary data., etc. But these methods may not be able to pick up other deeper aspects of laws that generate the return series. These deeper characteristics include nonlinearities, asymmetries, tail behaviors and other distribution characteristics that may be obscured and even twisted by model artifacts. For example, a conditional mean regression or time series model, the workhorse of financial analysis, is by design not able to look beyond the conditional mean of the returns data and its pattern and evolution. For highly volatile assets such as cryptocurrencies, but more generally for all non-Gaussian and nonlinear financial assets, these model-based comparisons and assessments are unnecessarily limiting. Assessment of "entire" distributions, made possible by entropy metrics, addresses these limitations. In addition, if the underlying laws are linear / Gaussian, nothing is lost. Entropy metrics become equivalent to these traditional conditional mean and variance models assessments.

In the first essay "Cryptocurrency Return Forecast Using Time Series Models and Entropy Approach", I investigate the notion of "similarity" between assets, in this case, cryptocurrencies with a set of other assets, including stock market indices. I use a novel approach based on entropy. This is in contrast to the traditional approach of comparing model fits, such as those based on time series models, including the GARCH model and ARIMA model. The approaches that based on time series models, which I also examine, reveal predictive structures and moments of underlying probability laws that generate returns. Entropies compare entire distributions of asset returns and capture all statistical aspects of returns and their differences. Both approaches are meant to identify new asset classes for diversification purposes. Sometimes assets are diverse in more ways, nonlinear and in higher moments and tails, than typical conditional mean and quantile models can reveal. Finally, I find max entropy closest industry portfolios to cryptocurrencies, which ensures shrinkage towards maximum diversification of portfolio weights. My findings will be useful in exploring the prediction of cryptocurrencies returns based on stock market performance.

In the second essay "Contrasting Cryptocurrencies with Other Assets: Full Distributions and the COVID Impact", I investigate any similarity and dependence based on the full distributions of cryptocurrency assets, stock indices and industry groups. I 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. I 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. I 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. I also find that the highest similarity before COVID-19 outbreak is between Bitcoin and Coal, Steel and Mining industries, and after COVID-19 outbreak it 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.

In the third essay “Do Cryptocurrencies and Other Assets Converge? A Clustering Analysis of Asset Returns”, I examine the prospects for clustering, or convergence of asset classes. In the first instance, I examine if a set of cryptocurrencies form identifiable clusters within this class. Using entropy metric to assess “similarity” of entire distributions, I implement Agglomerative Hierarchical Clustering technique to examine whether or not cryptocurrencies are converging to “clubs” with similar distributions of returns. To arrive at a more convincing conclusion, I also apply the K-means Clustering to justify our results. I discover that cryptocurrencies share similar geographic locations and similar functions tend to converge to same clusters. I also observe another potential explanation to our results called the “Coinbase effect”. In the second stage, I examine if these clusters include other asset classes, such as commodities. I find cryptocurrencies and commodities are separated into different clusters using entropy metric as cluster proximity, which is consistent with intuitive assumptions. I also find that the cluster that contains the distributions of Coal (COAL) and Petroleum and Natural Gas (OIL) have smaller distance to cryptocurrency distributions. To conclude, my work will help to enhance the profiling of the clusters with additional insights. As a result, this work offers a description of the market and a methodology that can be reproduced by investors that want to understand the main trends on the market and that look for cryptocurrencies with different financial performance.

All these three essays help to reveal the relationship between cryptocurrency returns and other asset returns. And I believe my findings will be useful in exploring the prediction of cryptocurrency returns based on stock market performance, and I verify that cryptocurrency is indeed an “orthogonal” assets that provide new opportunities to diversify risk.

Essays on Diversification: Cryptocurrencies and Other Assets

By

Xi Wu
M.S., Massachusetts Institute of Technology, 2015
B.S., Nankai University, 2014

Advisor: Esfandiar Maasoumi, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2022

Acknowledgments

Words cannot express my deepest thanks to dear advisor, Dr. Esfandiar (Essie) Maasoumi, who not only directed me to the final destination of this incredible doctoral journey, but also shaped me into a scholar equipped with established knowledge, critical thinking, and eagerness of exploring the world. During past years, Essie has been supportive to me in every regard: his smart ideas brought me into the magic world of Econometrics, his sharp insight illuminated my research direction and helped me out when I stuck during research, his generosity allowed me to embrace abundant resources for my research, his charming personality made him always a nice person to discussing and chatting with. It was my fortune to have him as my Ph.D. advisor and as a lifetime friend in the future.

My sincere appreciation also extend to Dr. Zhongjian Lin and Dr. Ruixuan Liu for being my dissertation committee members. Their guidance has been beyond just for this dissertation, and beneficial for me since the early stage of my Ph.D. study. Although it is not possible to list every name here, I appreciate the support from all of the faculty members, staff and friends in the Department of Economics, which I am grateful for being part of.

The very selfless persons who have supported me without any reservation are my parents. They respected my choice of studying abroad, despite the bitterness of letting the family be apart. Now, as I wrapping up my school life, there is a bigger world out there waiting for me to explore. I will be stepping onto the next chapter with my parents' best wishes. No matter how far I go, the inspiration from them is always planted in the deepest place of my heart.

On the road of my Ph.D. journey, I was not alone. I was super lucky to have my sweetest husband, Jeremy, who is also a Ph.D. student at Emory. I could not have these achievements without his love, support and company. I thank him for always seeing the best in me, bringing all the joy and happiness in my life. I look forward to

continuing our life journey together and wish all the best to his career path as well.

Contents

List of Figures

List of Tables

Chapter 1

Preface

This dissertation consists of three essays that study the financial market consists of cryptocurrencies and other traditional assets, like stock market indexes and commodities. In this Preface chapter, we will first answer the question that why we focus on cryptocurrency market, then demonstrate our motivation and contribution.

There are several things that investors do to protect their portfolios against risk. One significant way to protect one's portfolio is by diversifying. In short, this means an investor opts to include various types of asset classes, including bonds, stocks, commodities, REITs, cryptocurrencies., etc. The idea here is the same as the old adage "don't put your eggs all in one basket". When investors are invested in many areas, if one fails, the rest will ensure the portfolio as a whole remains secure. This added security can be measured in the increased profits that a diversified portfolio tends to bring in when compared to an individual investment of the same size. Therefore, it is key for investors to avoid choosing investments for their portfolios that are highly "similar". In this dissertation, we aim to study the "similarity" between cryptocurrencies and various traditional assets, thus provide guidance for portfolio diversification.

1.1 From Money and Fiat Currencies to Cryptocurrencies

Among a variety of emergent phenomena that we observe in human society, one of the most important is money. There are three functions of money: a store of value, a unit of account, and a medium of exchange. To facilitate trade by avoiding a problem of double coincidence of needs that restricts barter trading severely and inherently. According to economical models, a status of money is acquired in a process of the spontaneous symmetry breaking by a commodity that is the most easily marketable or, in other words, that is the most liquid one (Bak et al., 1999; Oswiecimka et al., 2013).

Fiat money is a type of money that is not backed by any commodity such as gold or silver, and typically declared by a decree from the government to be legal tender. Throughout history, fiat money was sometimes issued by local banks and other institutions. In modern times, fiat money is generally established by government regulation. Fiat money does not have intrinsic value and does not have use value. It has value only because the people who use it as a medium of exchange agree on its value. They trust that it will be accepted by merchants and other people (Goldberg, 2005).

The first cryptocurrency, Bitcoin, was proposed in 2008 (Nakamoto, 2008). The idea behind it was to decouple a currency from any institution or government, while preserving its status of a universal means of exchange, and to base a trust in this currency solely on a technology that supports it. Such a currency had to combine the advantages of both cash and electronic money: Anonymity of use (like cash) and capability of being transferred immediately to any place in the world (like electronic money). The already-existing technologies of asymmetric cryptography and distributed database (with a new consensus mechanism - “proof of work”) were linked into a decentralized secure register—blockchain that forms a staple of Bitcoin. Unlike traditional currencies, Bitcoin has inherently limited supply to prevent any loss of its value due to inflation (Wattenhofer, 2016).

1.2 Brief Introduction to Cryptocurrency Market

Cryptocurrency trading is possible, because they are easily convertible to traditional currencies like U.S. dollar (USD) or the Euro (EUR) and to other cryptocurrencies. This possibility is provided by 330 trading platforms (August 2020) open 24 hours a day, seven days a week. This, together with a fact that the most investors are individuals, distinguishes the cryptocurrency market from foreign exchange (Forex), where trading takes place from Monday to Friday essentially on the OTC market where mainly banks and other financial institutions participate in. Another peculiarity of the cryptocurrency market is that there is no reference exchange rate unlike Forex, where such reference rates are provided by Reuters. The sole exception is Bitcoin, whose exchange rate to USD is given by futures quoted on Chicago Mercantile Exchange (CME). Decentralization of the market means that the same cryptocurrency pairs are traded on different platforms, which—if accompanied by limited liquidity—can lead to sizeable valuation differences between platforms that produce arbitrage opportunities, both the dual and triangluar ones (Watorek et al., 2021; Makarov & Schoar, 2019; Gebarowski et al., 2019).

The spectacular development of a cryptocurrency market has attracted much interest of the scientific community. The first Bitcoin-related papers were published already in 2013–2015 (Kristoufek, 2013; Kristoufek, 2015), but a real boom on cryptocurrency-related publications occurred after 2017. Initially, only bitcoin was of significant interest (Bariviera et al., 2017; Drozdz et al., 2018; Garnier & Solna, 2019), but soon also other cryptocurrencies went under investigation (Wu et al., 2018; Kristoufek & Vosvrda, 2019). Then there appeared studies reporting on correlations within the market (Stosic et al., 2018; Bariviera at al., 2018; Bouri er al., 2019), and its relationship with regular markets (Corbet et al., 2018; Corelli, 2018). Recently, some researchers focused their attention on possible use of BTC as a hedging instrument for Forex (Urquhart & Zhang, 2019), for gold and other commodities (Shahzad et al., 2019), as well as for the stock markets (Shahzad et al., 2020; Wang et al., 2019).

The cryptocurrency market has already gone through a long route from a mere curiosity and a playground for the technology enthusiasts, via an emerging-market stage characterized by a relatively small capitalization, poor liquidity, large price fluctuations, short-term memory, frequent arbitrage opportunities, and weak complexity, to a more mature form characterized by medium capitalization, improved liquidity, inverse-cubic power-law fluctuations, long-term memory, sparse arbitrage opportunities, and increasing complexity.

The market of cryptocurrencies is fast and wild. Cryptocurrencies's philosophy is to break all borders and barriers, atleast associated with finance and trade. Since the inception of cryptocurrency, thousand of coins have been launched and are competing with each other even though it is in early stages of blockchain development. Every cryptocurrency which gets launched in the market comes with a unique promise that may turn the world around. In future there may be a single leader while others are rendered superseded, or there may be only 3-4 coins which will define the entire payments, lending, trading and banking infrastructures globally. It will be a new world, in a new light, in a new era (Thakur & Banik, 2018).

1.3 Motivation and Contribution of the Dissertation

The dissertation consists of three essays, which demonstrate the empirical applications of testing density similarity and nonlinear co-dependence with entropy metrics. We aim to find the density similarity between cryptocurrencies and traditional asset classes, thus to answer the question that if assets converge into any "clubs". The findings in this dissertation will offer a description of the market and a methodology that can be reproduced by investors that want to understand the main trends on the market and that look for cryptocurrencies with different financial performance.

In the first essay "Cryptocurrency Return Forecast Using Time Series Models and Entropy Approach", I investigate the notion of "similarity" between assets, in this case cryptocurrencies with a set of other assets, including stock market indices. I use a novel approach based on entropy. This is in contrast to the traditional approach of comparing model fit, such as those based on time series models, including GARCH model and ARIMA model. The time series model based approaches, which I also examine, reveal predictive structures and moments of underlying probability laws that generate returns. Entropies compare entire distributions of asset returns and capture all statistical aspects of returns, and their differences. Both approaches are meant to identify new asset classes for diversification purposes. Sometimes assets are diverse in more ways, nonlinear and in higher moments and tails, than typical conditional mean and quantile models can reveal. Finally, I find max entropy closest industry portfolios to cryptocurrencies, which ensures shrinkage towards maximum diversification of portfolio weights. My findings will be useful in exploring the prediction of cryptocurrencies returns based on stock market performance.

In the second essay "Contrasting Cryptocurrencies with Other Assets: Full Distributions and the COVID Impact", I investigate any similarity and dependence based on the full distributions of cryptocurrency assets, stock indices and industry groups. I 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. I 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. I 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. I 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.

In the third essay "Do Cryptocurrencies and Other Assets Converge? A Clustering Analysis of Asset Returns", I aim to examine the prospects for clustering, or convergence of asset classes. In the first instance, I examine if a set of cryptocurrencies form identifiable clusters within this class. Using entropy metric to assess "similarity" of entire distributions, I implement Agglomerative Hierarchical Clustering technique to examine whether or not cryptocurrencies are converging to "clubs" with similar distributions of returns. To arrive at a more convincing conclusion, I also apply the K-means Clustering to justify our results. I discover that cryptocurrencies share similar geographic locations and similar functions tend to converge to same clusters. I also observe another potential explanation to our results called the "Coinbase effect". In the second stage, I examine if these clusters include other asset classes, such as commodities. I find cryptocurrencies and commodities are separated into different clusters using entropy metric as cluster proximity, which is consistent with intuitive assumptions. I also find that the cluster that contains the distributions of Coal (COAL) and Petroleum and Natural Gas (OIL) have smaller distance to cryptocurrency distributions. To conclude, my work will help to enhance the profiling of the clusters with additional insights. As a result, this work offers a description of the market and a methodology that can be reproduced by investors that want to understand the main trends on the market and that look for cryptocurrencies with different financial performance.

To conclude, I make a central theme clear from the results in the essays: Diversification makes it desirable to identify new asset classes. The decision of what is a new asset class is based on its comparison with other assets. This comparison can be based on traditional methods which compare predictive models, such as GARCH and other time series models, summary data., etc. But these methods may not be able to pickup other deeper aspects of laws that generate the return series. These deeper characteristics include nonlinearities, asymmetries, tail behaviors and other distribution characteristics that may be obscured and even twisted by model artifacts. For a highly volatile assets such as cryptocurrencies, but more generally for all non Gaussian and nonlinear financial assets, these model based comparisons and assessments are unnecessarily limiting. Assessment of "entire" distributions, made possible by entropy metrics, addresses these limitations.

Bibliography

[1] Bak, Per, Simon F. Nørrelykke, and Martin Shubik. ”Dynamics of money.” Physical Review E 60, no. 3 (1999): 2528.

[2] Bariviera, Aurelio F., Luciano Zunino, and Osvaldo A. Rosso. ”An analysis of high-frequency cryptocurrencies prices dynamics using permutation-information-theory quantifiers.” Chaos: An Interdisciplinary Journal of Nonlinear Science 28, no. 7 (2018): 075511.

[3] Bariviera, Aurelio F., María José Basgall, Waldo Hasperué, and Marcelo Naiouf. ”Some stylized facts of the Bitcoin market.” Physica A: Statistical Mechanics and its Applications 484 (2017): 82-90.

[4] Berrett, Thomas B., Richard J. Samworth, and Ming Yuan. ”Efficient multivariate entropy estimation via K-nearest neighbour distances.” The Annals of Statistics 47, no. 1 (2019): 288-318.

[5] Bouri, Elie, Rangan Gupta, and David Roubaud. ”Herding behaviour in cryptocurrencies.” Finance Research Letters 29 (2019): 216-221.

[6] Corbet, Shaen, Andrew Meegan, Charles Larkin, Brian Lucey, and Larisa Yarovaya. ”Exploring the dynamic relationships between cryptocurrencies and other financial assets.” Economics Letters 165 (2018): 28-34.

[7] Corelli, Angelo. "Cryptocurrencies and exchange rates: A relationship and causality analysis." Risks 6, no. 4 (2018): 111.

[8] Drożdż, Stanisław, Jarosław Kwapień, Paweł Oświecimka, Tomasz Stanisz, and Marcin Watorek. "Complexity in economic and social systems: Cryptocurrency market at around COVID-19." Entropy 22, no. 9 (2020): 1043.

[9] Drożdż, Stanisław, Robert Gbarowski, Ludovico Minati, Paweł Oświcimka, and Marcin Watorek. "Bitcoin market route to maturity? Evidence from return fluctuations, temporal correlations and multiscaling effects." Chaos: An Interdisciplinary Journal of Nonlinear Science 28, no. 7 (2018): 071101.

[10] Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and H. Eugene Stanley. "A theory of power-law distributions in financial market fluctuations." Nature 423, no. 6937 (2003): 267-270.

[11] Garnier, Josselin, and Knut Solna. "Chaos and order in the bitcoin market." Physica A: Statistical Mechanics and its Applications 524 (2019): 708-721.

[12] Gebarowski, Robert, Paweł Oświecimka, Marcin Watorek, and Stanisław Drożdż. "Detecting correlations and triangular arbitrage opportunities in the Forex by means of multifractal detrended cross-correlations analysis." Nonlinear Dynamics 98, no. 3 (2019): 2349-2364.

[13] Goldberg, Dror. "Famous myths of" fiat money"." Journal of Money, Credit and Banking (2005): 957-967.

[14] Gopikrishnan, Parameswaran, Martin Meyer, LA Nunes Amaral, and H. Eugene Stanley. "Inverse cubic law for the distribution of stock price variations." The European Physical Journal B-Condensed Matter and Complex Systems 3, no. 2 (1998): 139-140.

[15] Kristoufek, Ladislav, and Miloslav Vosvrda. "Cryptocurrencies market efficiency ranking: Not so straightforward." Physica A: Statistical Mechanics and its Applications 531 (2019): 120853.

[16] Kristoufek, Ladislav. "BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era." Scientific reports 3, no. 1 (2013): 1-7.

[17] Kristoufek, Ladislav. "What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis." PloS one 10, no. 4 (2015): e0123923.

[18] Makarov, Igor, and Antoinette Schoar. "Price discovery in cryptocurrency markets." In AEA Papers and Proceedings, vol. 109, pp. 97-99. 2019.

[19] Nakamoto, Satoshi. "Bitcoin: A peer-to-peer electronic cash system." Decentralized Business Review (2008): 21260.

[20] Oświecimka, Paweł, Stanisław Drożdż, Robert Gebarowski, Andrzej Z. Górski, and Jarosław Kwapień. "Multiscaling edge effects in an agent-based money emergence model." arXiv preprint arXiv:1312.4803 (2013).

[21] Shahzad, Syed Jawad Hussain, Elie Bouri, David Roubaud, and Ladislav Kristoufek. "Safe haven, hedge and diversification for G7 stock markets: Gold versus bitcoin." Economic Modelling 87 (2020): 212-224.

[22] Shahzad, Syed Jawad Hussain, Elie Bouri, David Roubaud, Ladislav Kristoufek, and Brian Lucey. "Is Bitcoin a better safe-haven investment than gold and commodities?." International Review of Financial Analysis 63 (2019): 322-330.

[23] Stosic, Darko, Dusan Stosic, Teresa B. Ludermir, and Tatijana Stosic. "Collective behavior of cryptocurrency price changes." Physica A: Statistical Mechanics and its Applications 507 (2018): 499-509.

[24] Thakur, Kumar Krishnan, and G. G. Banik. ”Cryptocurrency: Its Risks and Gains and The Way Ahead.” IOSR Journal of Economics and Finance 9, no. 2 (2018): 38-42.

[25] Urquhart, Andrew, and Hanxiong Zhang. ”Is Bitcoin a hedge or safe haven for currencies? An intraday analysis.” International Review of Financial Analysis 63 (2019): 49-57.

[26] Wang, Pengfei, Wei Zhang, Xiao Li, and Dehua Shen. ”Is cryptocurrency a hedge or a safe haven for international indices? A comprehensive and dynamic perspective.” Finance Research Letters 31 (2019): 1-18.

[27] Watorek, Marcin, Stanisław Drożdż, Jarosław Kwapień, Ludovico Minati, Paweł Oświecimka, and Marek Stanuszek. ”Multiscale characteristics of the emerging global cryptocurrency market.” Physics Reports 901 (2021): 1-82.

[28] Wattenhofer, Roger. The science of the blockchain. Inverted Forest Publishing, 2016. Wu, Ke, Spencer Wheatley, and Didier Sornette. ”Classification of cryptocurrency coins and tokens by the dynamics of their market capitalizations.” Royal Society open science 5, no. 9 (2018): 180381.

Chapter 2

Cryptocurrency Return Forecast Using Time Series Models and Entropy Approach

In this chapter, we investigate the notion of "similarity" between assets, in this case cryptocurrencies with a set of other assets, including stock market indices. We use a novel approach based on entropy. This is in contrast to the traditional approach of comparing model fit, such as those based on time series models, including GARCH model and ARIMA model. The time series model based approaches, which we also examine, reveal predictive structures and moments of underlying probability laws that generate returns. Entropies compare entire distributions of asset returns and capture all statistical aspects of returns, and their differences. Both approaches are meant to identify new asset classes for diversification purposes. Sometimes assets are diverse in more ways, nonlinear and in higher moments and tails, than typical conditional mean and quantile models can reveal. Finally, we find max entropy closest industry portfolios to cryptocurrencies, which ensures shrinkage towards maximum diversification of portfolio weights. Our findings will be useful in exploring the prediction of cryptocurrencies returns based on stock market performance.

Keywords:

Cryptocurrency, Bitcoin, Entropy, GARCH, Optimal Portfolio

2.1 Introduction

In recent years, financial markets witnessed the birth and development of a new assets class, the cryptocurrency, and it is receiving significant attention. On the one hand it is based on a fundamentally new technology, the potential of which is not fully understood. On the other hand, at least in the current form, it fulfills similar functions as other more traditional assets.

The starting point of the development of cryptocurrency was in the year of 2008, when the Bitcoin emerged, based on blockchain technology. The cryptocurrency market is an important markets in the global assets markets. As of February 2019, there were over 17.53 million bitcoins in circulation with a total market value of around $63 billion. Today, there are literally thousands of cryptocurrencies in existence, with an aggregate market value of over $120 billion.

With the rapid development of cryptocurrency market, the literature has focused on statistical properties and risk behavior of the cryptocurrency by comparing them with classical assets like equities and exchange rates. 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), Baur et al. (2018), Stavroyiannis (2018) and Catania and Grassi (2017) observed the phenomenon of volatility clustering in cryptocurrency market. Osterrieder and Lorenz (2017) and Begusic et al. (2018) 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. Regime-switching behaviors are detected by Bariviera et al. (2017), Balcombe and Fraser (2017). Thies and Molnar (2018) have identified structural breaks in the volatility process of bitcoin via a Bayesian framework. Recently, Lahmiri et al. (2018) and Lahmiri and Bekiros (2018) have pointed out that Bitcoin markets are characterized by long memory and multi-fractality.

Most of the existing studies focus on Bitcoin returns. 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 discuss that if Bitcoin belongs to any 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 between currency and commodity (Dyhrberg, 2016).

Another area evoke people's interest is forecasting Bitcoin volatility, because 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.

GARCH-type models have been employed to forecast cryptocurrency market volatility. Dyhrberg (2016) explores Bitcoin volatility using GARCH models and suggests that Bitcoin has several similarities with both gold and the dollar. Katsiampa (2017) explores the applicability of several ARCH-type specifications to model Bitcoin volatility and selects an AR-CGARCH model as the preferred specification. Conrad et al. (2018) used the GARCH-MIDAS model to extract the long- and short-term volatility components of cryptocurrencies. Segnon and Bekiros (2019) found that the Markov switching multifractal (MSM) and FIGARCH models outperform other GARCH-type models in forecasting Bitcoin returns volatility.

Our objective in this chapter is to revisit some stylized facts of cryptocurrency markets and propose new econometrics models that produce accurate volatility forecasts. In contrast to previous studies that use time series models to forecast cryptocurrency returns, in this chapter we also use entropy profiles method to test the density similarity between cryptocurrency and stock returns. We not only consider Bitcoin as most of the literatures did, but also consider Ethereum, since both of them are leading cryptocurrency markets which have large volume and long history. We use nonparametric entropy metrics to test equality between cryptocurrency density and stock market index density. Enytropy metrics outperform time series models (ARIMA and GARCH) in testing the equality between two densities without knowing the exact form of the distribution, while time series models have advantage in that they can tell the model revolution. Another innovative method of this chapter is that we propose a maximum entropy approach to optimal portfolio selection, which let us to construct a portfolio with stocks in three industries (Coal, Steel and Mines) which have most similar densities with Bitcoin.

The rest of the chapter is organized as follows. Section 2 presents the data analysis and some stylized facts. In Section 3, we provide the statistical properties of our proposed models and study in detail their forecasting performance and adequacy by time series models (ARIMA model and GARCH model). Also, we display the empirical results in terms of graphs and regression results. In Section 4, we calculate nonparametric entropy metrics to test the density equality between two cryptocurrencies (Bitcoin and Ethereum) and two stock market indexes (SP500 and NASDAQ). We conduct equality tests on both marginal distributions and conditional distributions. We find that NASDAQ has the most similar density with Bitcoin. Finally, we revisit the current mean-variance portfolio optimization method, and propose a maximum entropy approach to portfolio selection, which ensures shrinkage towards maximum diversification of portfolio weights. Section 5 provides the concluding remarks.

2.2 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) range from July 16, 2010 to April 14, 2019; the price observations of Ethereum (ETH-USD) range from August 6, 2015 to April 14, 2019; the price observations of S&P500 stock market index (ˆGSPC) range from January 1, 2010 to April 12, 2019; the price observations of NASDAQ stock market index (ˆIXIC) range from January 1, 2010 to April 14, 2019. In each data set of cryptocurrency 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 cryptocurrency market data and stock market indexes, we calculate the daily log return using adjusted close price:

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

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 historical highest price in December 2017. Bitcoin's price rose 5% in 24 hours, with its value being up 1,824% since 1 January 2017, to reach a new all-time high on December 17, 2017. After this price peak, the cryprocurrency price dropped

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

dramatically. The descriptive statistics of daily log-returns are reported in Table 1. The daily returns of cryptocurrency markets exhibit high variability and excess kurtosis. These deviations from the Normal distribution, as in Figure 2, are confirmed by the Jarque-Bera test that rejects the null hypothesis of normality. Figure 2 illustrates the unconditional distributions of Bitcoin daily returns and Ethereum daily returns, we observe that the unconditional distributions of Bitcoin returns and Ethereum returns do not converge to Normal distribution, as shown in Figure 2, and this is also consistent with our Jarque-Bera test result.

We applied the Augmented-Dicker-Fuller (ADF) unit-root test of Dickey and Fuller (1979), 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-stationarity. 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 cryptocurrency market is not as efficient as stock or foreign exchange markets, which display a complete lack of predictability (Lahmiri et al., 2018).

Table 2.1: Descriptive statistics

Daily log-return	S&P500	NASDAQ	Bitcoin	Ethereum
Observations	2335	2335	3192	1345
Mean	0.04	0.05	0.36	0.30
Standard deviation	0.94	1.08	6.75	7.55
Skewness	-0.47	-0.44	2.96	-1.13
Kurtosis	4.60	3.48	93.10	18.83
Augmented Dickey-Fuller	-14.16	-14.13	-13.49	-9.75
Jarque-Bera	2147.55	1254.50	1158680	20216.24

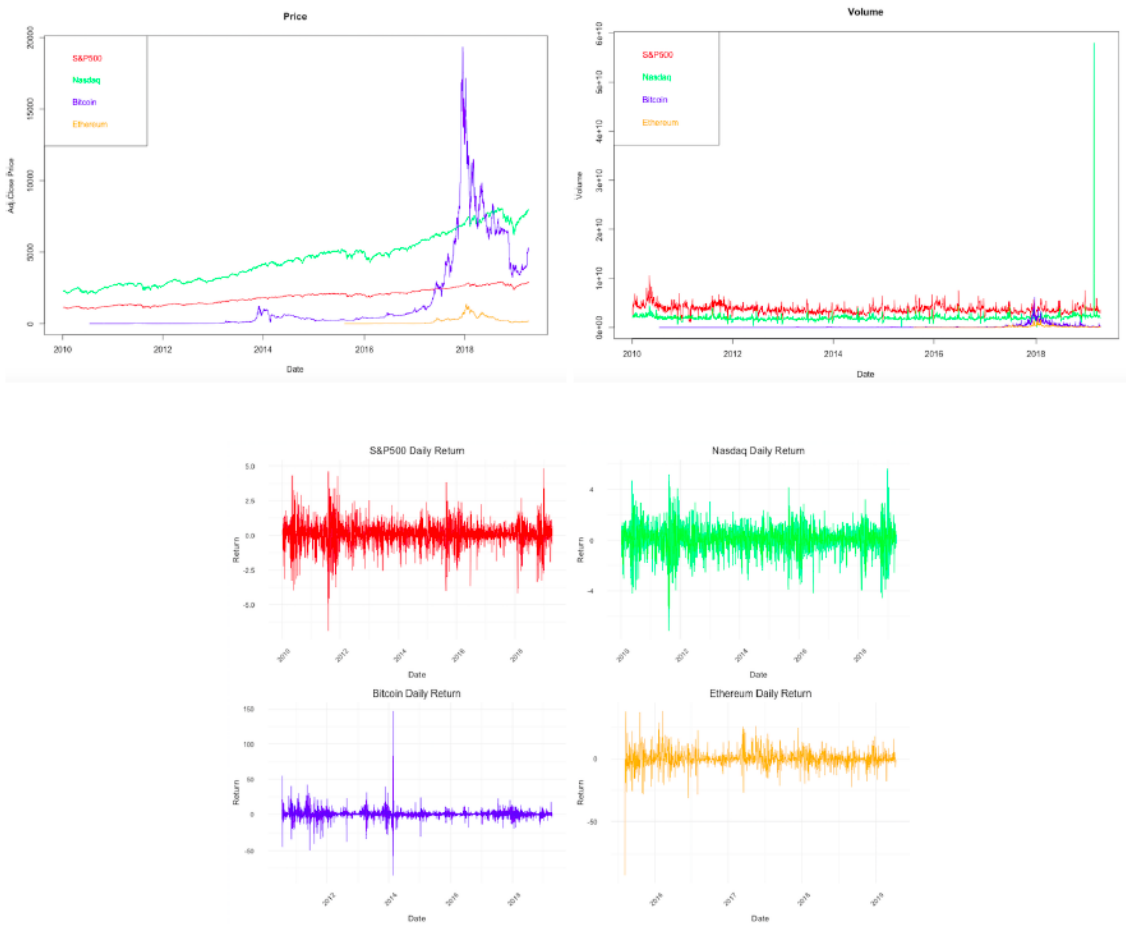


Figure 2.1: Plot of price, volume and daily log-returns

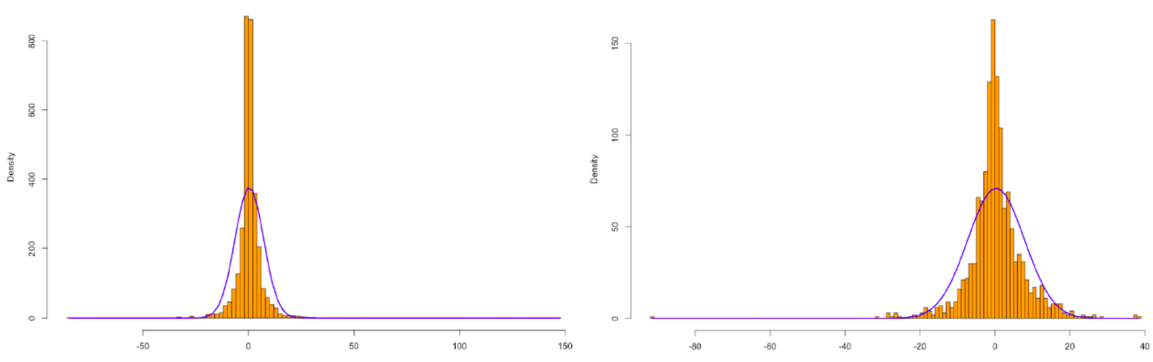


Figure 2.2: Plot of Bitcoin & Ethereum returns distributions

2.3 Time Series Method

In this part, we use two time series models (ARIMA model and GARCH model) to forecast cryptocurrency returns. We choose to apply time series models because: First, the best of the ARIMA models is used to model the linear data of time series and the residual of this linear model will contain only the nonlinear data. The ARIMA model specifies the conditional mean of the process. Second, the GARCH is used to model the nonlinear patterns of the residuals, it specifies the conditional variance of the process. We use these two models to analyze the univariate series and to predict the values of approximation. The time series models are useful because they can give exact forms of the model, in other words, we can get a model with estimated parameters as a result.

2.3.1 ARIMA Model

We assume that returns $\{r_t\}$ in cryptocurrency markets follow autoregressive fractionally integrated moving average process (ARIMA) (Granger and Joyeux, 1980; Hosking, 1981), given by the following equation:

$$\Phi(L)(1-L)^d(r_t-\mu)=\Theta(L)\epsilon_t.$$

The lag polynomials are defined as:

$$\Phi(L)=1-\phi_1 L-\phi_2 L^2-...-\phi_p L^p,$$

$$\Theta(L)=1-\theta_1 L-\theta_2 L^2-...-\phi_q L^q,$$

where the p and q are autoregressive (AR) and moving average (MA) orders, respectively. L is the lag operator. The innovation process, ϵ_t can be formalized as:

$$\epsilon_t = u_t \sigma_t,$$

where u_t is a sequence of independent identically distributed normal random variables with zero mean and unit variance.

The ARIMA model is suggest for the nonstationary and nonseasonality data series. ARIMA model is a type of models in the Box-Jenkins modeling. The Box-Jenkins methodology includes four iterative steps of model identification, parameter estimation, diagnostic checking and forecasting. In identification step, data transformation is required to make the series stationary. The stationary process is a necessary condition in building an ARIMA model. When the observed time series presents trends and nonseasonal behavior, data transformation and differencing are applied to the data series in order to stabilize variance and to remove the trend before an ARIMA model is applied. In the Box-Jenkins modeling, the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the sample data are used in identifying the order of the time series model. The chosen model then is statistically checked whether it accurately describes the series. The model fits well if the P-value of its parameter is statistically significant, as well as its residuals are generally small, randomly distributed, and contain no useful information, where at this point, the model can be used for forecasting.

We use maximum likelihood estimation (MLE) to estimate the ARIMA model. The MLE method tries to maximize the log-likelihood for given values of p, d, and q when finding parameter estimates so as to maximize the probability of obtaining the data that we have observed. The p,d, and q are chosen by minimizing the AIC. The algorithm uses a stepwise search to traverse the model space to select the best model with smallest AIC.

We use correlogram analysis to determine p and q in ARIMA model, the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the sample data are used in identifying the order of the time series model.

The fitting result for Bitcoin return is an ARIMA(5,1,0) process, and the fitting result for Ethereum return is an ARIMA(1,0,0) process. The correlogram analysis (ACF and PACF) for Bitcoin and Ethereum daily log-returns are shown in Figure 3. The coefficients of the ARIMA process are shown in Table 2. From the result, we can see the models are not correctly specified, that will usually be reflected in residuals in the form of trends, skeweness, or any other patterns not captured by the model. So, ARIMA model cannot forecast the daily log-returns of cryptocurrency market perfectly.

Table 2.2: ARIMA model result

ARIMA	Bitcoin–ARIMA(5,1,0)	Ethereum–ARIMA(1,0,0)
Intercept	-	0.2992
AR1	-0.8372	-0.0587
AR2	-0.8465	-
AR3	-0.6398	-
AR4	-0.3741	-
AR5	-0.1275	-
AIC	21518.29	9249.24
BIC	21554.7	9264.86

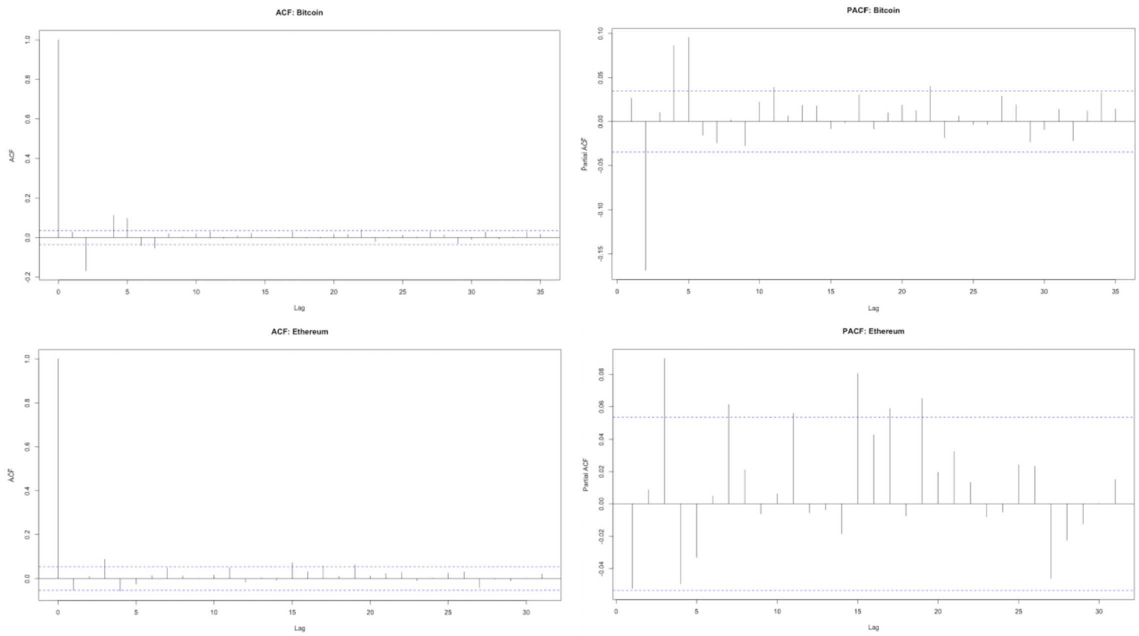


Figure 2.3: Correlogram analysis

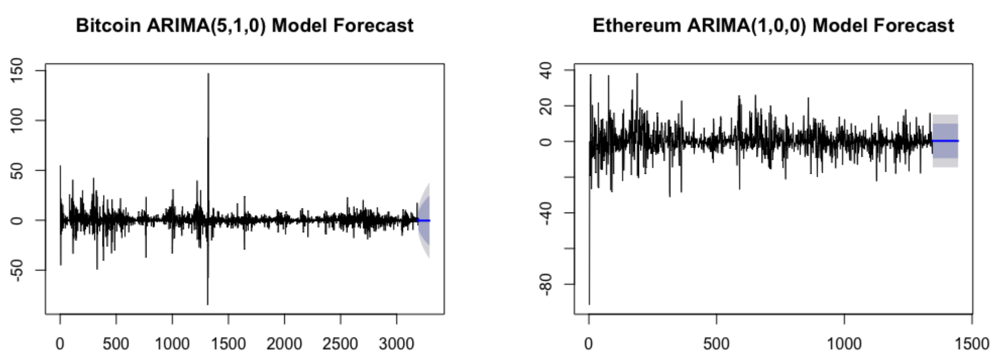


Figure 2.4: Forecast using ARIMA models

2.3.2 GARCH Model

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) is an extension of the ARCH model that incorporates a moving average component together with the autoregressive component. In this chapter, we consider the mean model of AR(1). The standard GARCH(p,q) model can be written as:

$$y_t = x'(t)b + \epsilon_t,$$

$$\epsilon_t|\Psi_{t-1} \sim N(0, \sigma_t^2),$$

$$\sigma_t^2 = (\omega + \sum_{j=1}^{m} \zeta_j \nu_{jt}) + \sum_{j=1}^{q} \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2,$$

where σ_t^2 denoting the conditional variance, ω the intercept and ϵ_t^2 the residuals from the mean filtration process.

The GARCH order is defined by (p, q) (ARCH,GARCH), where p is order of GARCH terms σ^2 and q is the ARCH terms ϵ^2, with possibly m external regressors ν_j which are passed pre-lagged. If variance targeting is used, then ω is replaced by,

$$\tilde{\sigma}^2(1 - \hat{P}) - \sum_{j=1}^{m} \zeta_j \bar{\nu_j}$$

where $\tilde{\sigma}^2$ is the unconditional variance of ϵ^2 which is consistently estimated by its sample counterpart at every iteration of the solver following the mean equation filtration, and $\bar{\nu_j}$ represents the sample mean of the j^{th} external regressors in the variance equation, and $\hat{P}$ is the persistence. One of the key features of the observed behavior of financial data which GARCH models capture if volatility clustering which may be quantified in the persistence parameter $\hat{P}$, which can be defined below,

$$\hat{P} = \sum_{j=1}^{q} \alpha_j + \sum_{j=1}^{p} \beta_j.$$

Finally, the unconditional variance of the model $\hat{\sigma}^2$, and related to its persistence, is,

$$\hat{\sigma}^2 = \frac{\hat{\omega}}{1 - \hat{P}},$$

where $\hat{\omega}$ is the estimated value of the intercept from the GARCH model.

In Table 3, we get the estimated parameters of the GARCH model for daily returns of S&P500, NASDAQ, Bitcoin and Ethereum market, respectively. The $\hat{\alpha_1}$ is the coefficients to the squared lag residuals in the GARCH equation, and $\hat{\beta_1}$ is the coefficients to the lag variance. The GARCH parameters can be interpreted that the large value of $\hat{\beta_1}$ cause σ_t to be highly correlated with σ_{t-1} and gives the conditional standard deviation process a relatively long-term persistence, at least compared to its behavior under an ARCH model. We can see from Table 3 that all the $\hat{\omega}$ in the three marktets (NASDAQ, Bitcoin and Ethereum) are statistically significant, implying that this is a small amount of positive autocorrelation. Also, all the $\hat{\beta_1}$ and $\hat{\alpha_1}$ are highly significant, which implies rather persistent volatility clustering.

In Figure 4, we have the GARCH patterns which plot the trend of estimated conditional variances (green lines) and the squared residuals (black lines). We can conclude from the figure that the estimated conditional variances and squared residuals have the same trend, so the estimated conditional variances perfectly depict the squared residuals of series. The conditional variance processes are second moment terms, similarly that for the mean process, we are able to estimate the unconditional variance and residuals in GARCH(1,1):

$$\sigma_t^2 = a + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2;$$

$$r_t - \mu_t = \epsilon_t = \sigma_t z_t.$$

where z_t is an i.i.d. process with $E_t(z_t) = 0$ and $Var_t(z_t) = 1$. Once we fit our conditional variance models we will be left with the conditional variance process σ_t^2. At this point we know the conditional variance process σ_t^2 and ϵ_t^2. This allow us to obtain the final standardized residuals series z_t which is i.i.d and equal to $\epsilon_t/\sigma_t = z_t$. We can see that in a S&P500, NASDAQ, Bitcoin and Ethereum markets, the estimated conditional variances and estimated squared residuals move in the same direction. This is coordinated with our assumptions of GARCH(1,1) model.

In Figure 5, we can see how the forecast of the conditional variance picks up from the last estimated conditional variance. The black lines are the last 20 residuals, while the green lines are conditional variances forecast by GARCH, and the orange lines are the unconditional forecast continuous from the conditional forecast. In fact, we can see the conditional variance decreases from last estimated ones, slowly, towards the unconditional variance value, which indicate the accuracy of GARCH model forecast.

Table 2.3: GARCH Estimation

Daily log-return	Parameter	Estimate	Std.Error	t Value	Pr(>t)
S&P500	$\hat{\omega}$	0.035544	0.005596	6.3520	0.00000
	$\hat{\alpha_1}$	0.161659	0.018509	8.7340	0.00000
	$\hat{\beta}_1$	0.800386	0.019371	41.3186	0.00000
NASDAQ	$\hat{\omega}$	0.051727	0.008963	5.7714	0.00000
	$\hat{\alpha_1}$	0.127917	0.016000	7.9948	0.00000
	$\hat{\beta}_1$	0.826103	0.019636	42.0717	0.00000
Bitcoin	$\hat{\omega}$	0.774732	0.102623	7.5493	0.000000
	$\hat{\alpha_1}$	0.207986	0.015613	13.3210	0.000000
	$\hat{\beta}_1$	0.791014	0.013375	59.1411	0.000000
Ethereum	$\hat{\omega}$	3.677801	0.758113	4.8513	0.000001
	$\hat{\alpha_1}$	0.189519	0.027947	6.7814	0.000000
	$\hat{\beta}_1$	0.742106	0.032977	22.5039	0.000000

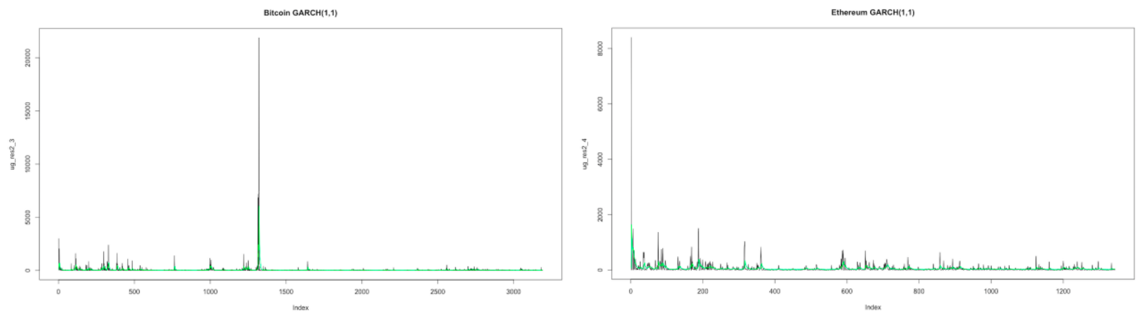


Figure 2.5: GARCH pattern

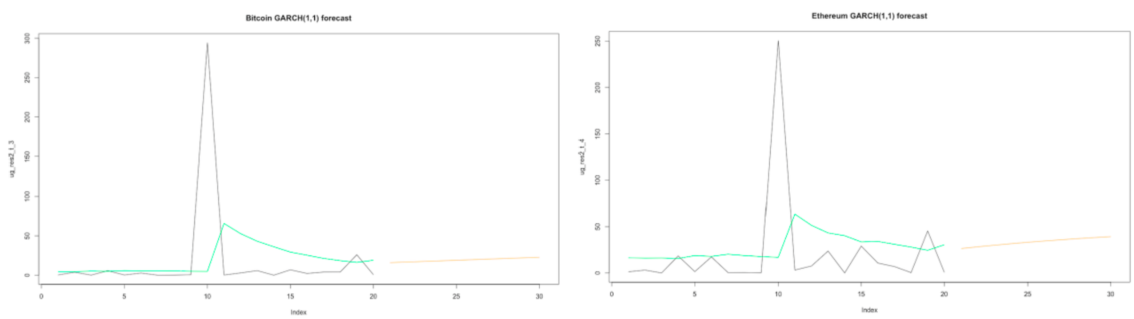


Figure 2.6: GARCH forecast

In this part, we investigtated GARCH model because it is viewed as attempts to capture some nonlinearities. Hong & White (2000) removed persistent GARCH effects from the S&P500 series. In other financial applications, the presence of unaccounted nonlinear dependence in the residuals of such models has been detected. Pagan & Schwert (1990) studied the performance of parametric, nonparametric, as well as semiparametric models of conditional variances in the GARCH setting for monthly U.S. stock returns. Notably, in most studies the burden of dealing with nonlinearities is placed on the successful modeling of the conditional mean and conditional variance. A common conclusion from these studies is that the observed nonlinearities may be too complex to be exploitable for improved predictability, especially with small samples

(Stengos, 1995; Hsieh, 1989).

However, GARCH model still has limitations in forecasting the conditional variance of cryptocurrency return. One of the well established facts in the financial modeling states that empirical distributions of log-returns time series are skewed and fat-tailed. Therefore, there are various models which have been introduced that resort to an alternative non-Gaussian assumption. Since the skewness defines the asymmetry of the distribution, it has a significant impact on the shape of the tails. Kurtosis is another parameter of interest under the alternative non-Gaussian fat-tailed assumption. As a result, it is crucial to model skewness and kurtosis as accurately as possible, also considering changes over time. Given the sensitivity of parametric GARCH to misspecification of the mean and variances, its nonparametric implementations are worthy of further research.

2.4 Entropy Profiles Method

2.4.1 Introduction to Information Theory and Entropy

Let us suppose we observe two threshold traits X and Y which are possibly correlated. This correlation referring to corresponding for X and Y liabilities cannot be measured by Pearson' correlation coefficient because the values of X and Y are not observable in the continuous scale. So, we need to use a measure of correlation for the categorical values of X and Y, which is the entropy. In this part of the chapter, we will give a brief introduction to information theory and entropy.

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, within the context of communication process, to develop his information criterion. His criterion, called Shannon's entropy, is

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

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

$$H(x) = -\int p(x)log(p(x))dx,$$

where this differential entropy does not have all of the properties of the discrete entropy. The Shannon's entropy has three main properties:

1. $H(x) = 0$ if and only if when there exist one event x with $p(x) = 1$;

2. The value of entropy reaches the maximum when all events x have the same probability;

3. For two independent variables X and Y, $H(x, y) = H(x) + H(y)$.

After introducing the Shannon's entropy measure, a fundamental question arose: whose information does this measure capture? what does it measure? One answer to this question is that H is a measure of the amount of information in a message. To measure information, one must abstract away from any form or content of the message itself. Further, 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 claim that the outcome of an experiment contains the amount of information $H(p)$.

The information discussed here is not subjective information of a particular researcher. The information observed in a single observation, or a data set, is a certain quantity that is independent of whether the observer recognizes it or not. Thus, $H(p)$ is a measure of the average amount of information provided by an outcome of a random drawing governed by p. 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.

According to both Shannon and Jaynes, H measures the degree of ignorance of a communication engineer who designs the technical equipment of a communication channel because it takes into account the set of all possible messages to be transmitted over this channel during its lifetime.

In addition, some prior information q, defined on $\Re$, exists, the cross-entropy (Kullback-Leibler, K–L, 1951) measure is

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

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.

Facing the fundamental question of drawing inferences from limited and insufficient data, Jaynes (1982) proposed the maximum entropy (ME) principle, which he viewed as a generalization of Bernoulli and Laplace's Principle of Insufficient Reason.

Given T structural constraints in the form of moments of the data (distribution), Jaynes proposed the ME method, which is to maximize $H(p)$ subject to the T structural constraints. Thus, if we have partial information in the form of some 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

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

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 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

$$H_\alpha^R(p) = \frac{1}{1-\alpha} log \sum_k p_k^\alpha.$$

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

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

which is equal to the traditional cross-entropy measure as $\alpha \to 1$.

Metric entropy applied to finance theory has been successfully developed mainly after 1999. According to Gulko (1999), the "entropy pricing theory" suggests that in informational efficient markets, perfectly uncertain market beliefs must prevail. When the entropy functional is used to index the market uncertainty, then the entropy-maximizing market beliefs must prevail. To optimize various entropic measures, one can derive new asset pricing models that are similar to Black-Scholes model with the log-normal distribution replaced by other probability distributions.

Entropic methodology was also widely applied within computer science and machine learning fields. Many statistical procedures, including goodness-of-fit tests and methods for independent component analysis, rely critically on the estimation of the entropy of a distribution. Scholars seek entropy estimators that are efficient and achieve the local asymptotic minimax lower bound with respect to squared error loss (Berrett et al., 2019). These results facilitate the construction of asymptotically valid confidence intervals for the entropy of asymptotically minimal width.

2.4.2 Using Entropy to Test Equality of Univariate Densities

The reason why we prefer to use metric entropy to reveal the similarity structure between cryptocurrency and stock market is to solve the misspecification problem by ARIMA and GARCH model. We would like to make an inquiry about any unconditional or conditional similarity structure without requiring the specification of conditional Mean-Variance models. Therefore, we will focus our research on nonparametric density and other functional estimation. Entropies are defined directly in terms of the actual distributions and not the variables and their moments. Partly for this reason they also offer a clearer view of the relation between total independence, conditional similarity, and causality relations in several directions.

Accordingly, we propose to study the cryptocurrency and stock returns for unconditional, nonparametric similarity using the Kullback-Leibler (KL) (1951) measure which we have found to be successful in detecting generic and possibly nonlinear similarity (Granger et al., 2000). The KL measure are seeing increasing and welcome use in testing for independence and other hypothesis. For instance, Robinson (1991), Delgado (1994), Hong & White (2000), and Zheng (2000) are all concerned with the KL measure for testing independence. Being entropy based, this measure is defined over the densities of the cryptocurrency and stock returns which we estimate nonparametrically.

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 of Granger & Maasoumi (2004)) for testing the null of equality of two univariate density (or probability) functions. For continuous variables,

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

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 cryptocurrency density and stock market index density. The following properties are satisfied by this entropy measure S_ρ (Granger et al., 2000):

1. S_ρ is well defined for both continuous and discrete variables;

2. S_ρ is normalized to zero if X and Y are independent, and lies between 0 and 1;

3. The modulus of the entropy measure S_ρ is equal to unity if there is a measurable exact (nonlinear) relationship, $Y = g(X)$ say, between the random variables;

4. S_ρ is equal to or has a simple relationship with the (linear) correlation coefficient in the case of a bivariate normal distribution;

5. S_ρ is metric, that is, it is a true measure of distance and not just of Kullback-Leibler divergence;

6. The entropy measure S_ρ is invariant under continuous and strictly increasing transformations $h(\cdot)$. This is useful since X and Y are independent if and only if

$h(X)$ and $h(Y)$ are independent.

Granger et al. (2000) consider a kernel implementation of this metric and demonstrate how critical values can be obtained under the null of univariate densities equality in the case of time series data. For the estimation of the univariate densities in S_ρ we use kernel density estimators. For the kernel function we employ the widely used nonparametric second-order Gaussian kernel, while bandwidths are selected via likelihood cross-validation (Silverman, 2018). The block bootstrap is conducted via resampling with replacement from the pooled empirical distributions of X and Y.

We apply the metric S_ρ to the daily returns data by setting $x = Return_{cryptocurrency}$ and $y = Return_{stock}$. Table 4 shows the result of the S_ρ value and corresponding p-value. 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 independence to detect significant deviation from zero.

Examining Table 4 we see that value of S_ρ is smaller when $x =$ Bitcoin and $y =$ NASDAQ, 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 6 and Figure 7, we can also see the Bitcoin daily returns density and the NASDAQ stock market index daily returns density have similar shapes. Therefore, it would be more meaningful for us to study the similarity structure between Bitcoin and NASDAQ markets.

Table 2.4: Consistent univariate entropy density equality test

	S_rho	p-value
S&P500 & Bitcoin	0.2003	<2.22e-16
S&P500 & Ethereum	0.3273	<2.22e-16
NASDAQ & Bitcoin	0.1705	<2.22e-16
NASDAQ & Ethereum	0.2919	<2.22e-16

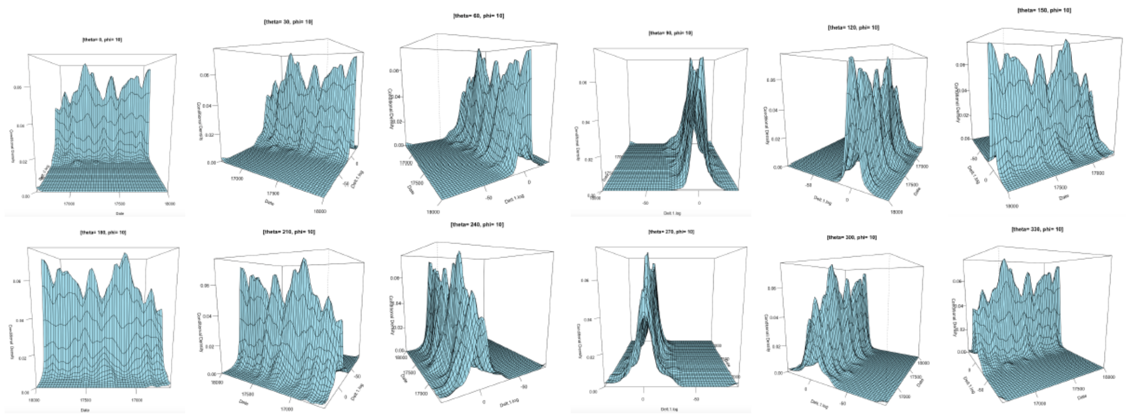


Figure 2.7: Density of NASDAQ

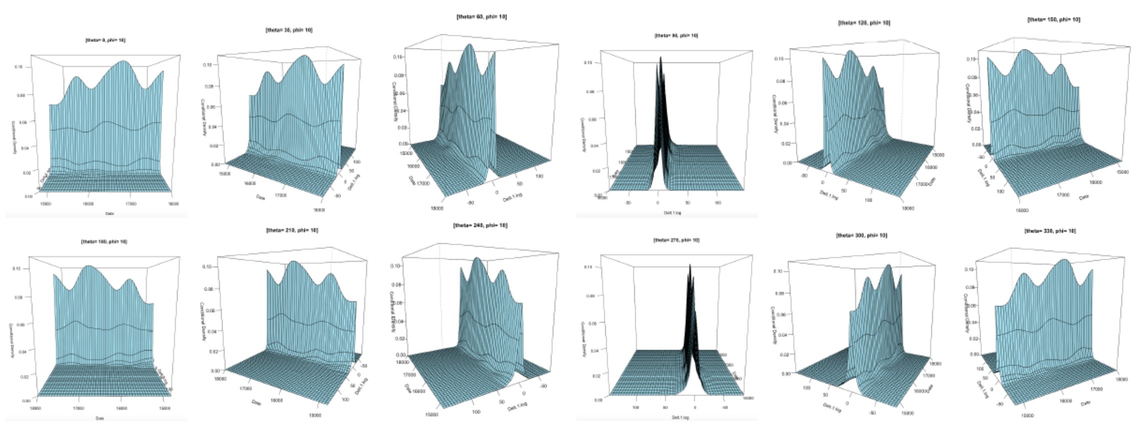
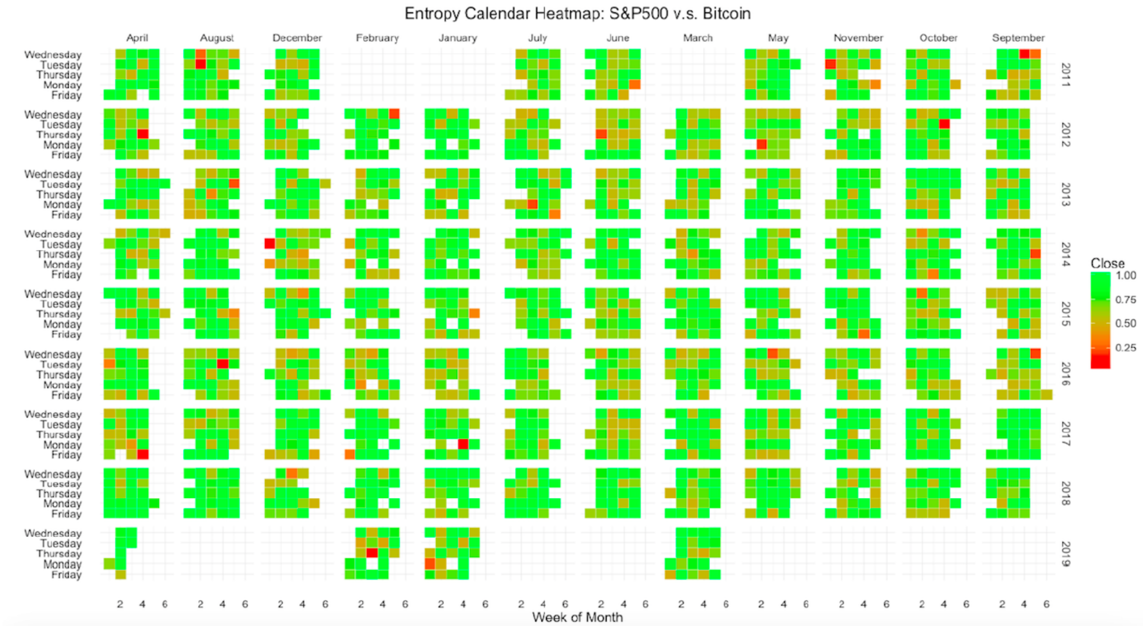


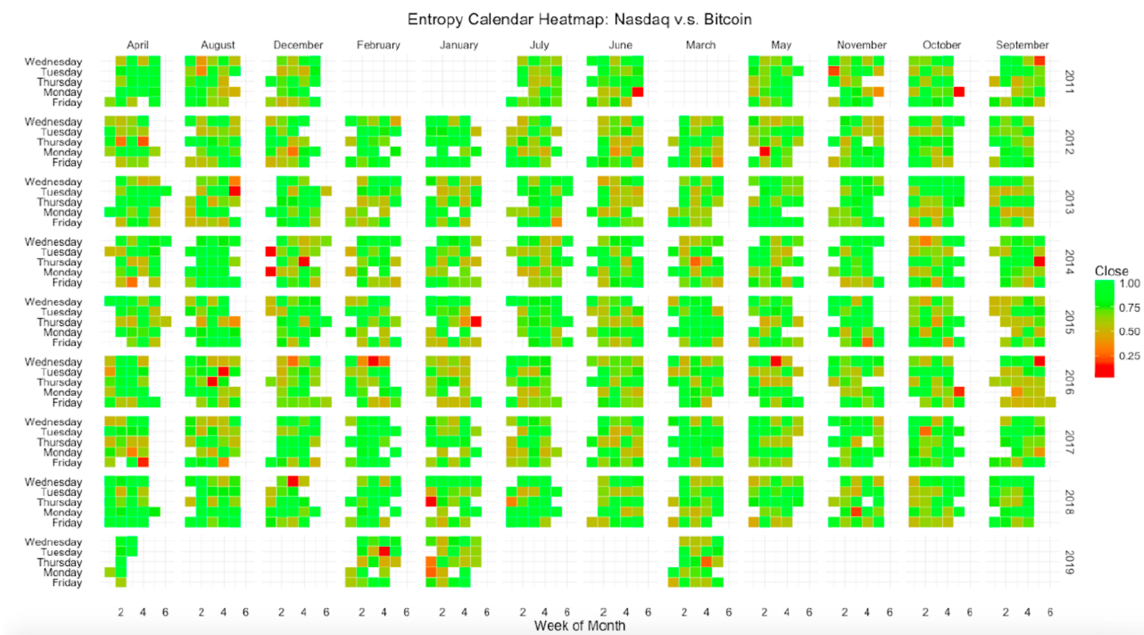
Figure 2.8: Density of Bitcoin

To learn the calendar effect of similarity structure between Bitcoin and NASDAQ return densities, we analyze the variation of entropy measure S_ρ on different weekdays and month. Table 5 shows the value of S_ρ entropy measure and corresponding p-value on different months as well as different weekdays. We conduct this calendar effect study because investors want to see the variation, especially the highs and lows of a metric on an actual calendar itself, since it emphasizes the variation over time rather than the actual value itself. Figure 8 shows the movement of S_ρ entropy measure along months and weekdays. The time period is from the year of July 19, 2010 to the April 10, 2019. In each year we have 12 trading months, and in each week we have 5 trading days (Monday to Friday). In contrast to the stocks, there is no pronounced Monday effect for Bitcoin. From the result we can see the S_ρ entropy measures are lower in September, which is 0.1262. This indicates that the distance between Bitcoin return density and NASDAQ return density is smaller in September than other months. The S_ρ entropy measures are quite stable along the weekdays, so we will not discuss the economic interpretation about this.

Table 2.5: Calendar effect of entropy measure

	Jan	Feb	Mar	Apr	May	Jun
S_rho	0.2194	0.2509	0.2043	0.2437	0.2315	0.1849
p-value	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16
	Jul	Aug	Sep	Oct	Nov	Dec
S_rho	0.2151	0.1746	0.1262	0.1645	0.2452	0.2429
p-value	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16
	Mon	Tue	Wed	Thu	Fri	
S_rho	0.2159	0.1850	0.1931	0.2080	0.1763	
p-value	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	


Entropy Calendar Heatmap: S&P500 v.s. Bitcoin
April
August
December
February
January
July
June
March
May
November
October
September
Wednesday
Tuesday
Thursday
Monday
Friday
2011
2012
2013
2014
2015
2016
2017
2018
2019
Close
1.00
0.75
0.50
0.25
2 4 6
Week of Month


Entropy Calendar Heatmap: Nasdaq v.s. Bitcoin
April
August
December
February
January
July
June
March
May
November
October
September
Wednesday
Tuesday
Thursday
Monday
Friday
2011
2012
2013
2014
2015
2016
2017
2018
2019
Close
1.00
0.75
0.50
0.25
2 4 6
Week of Month

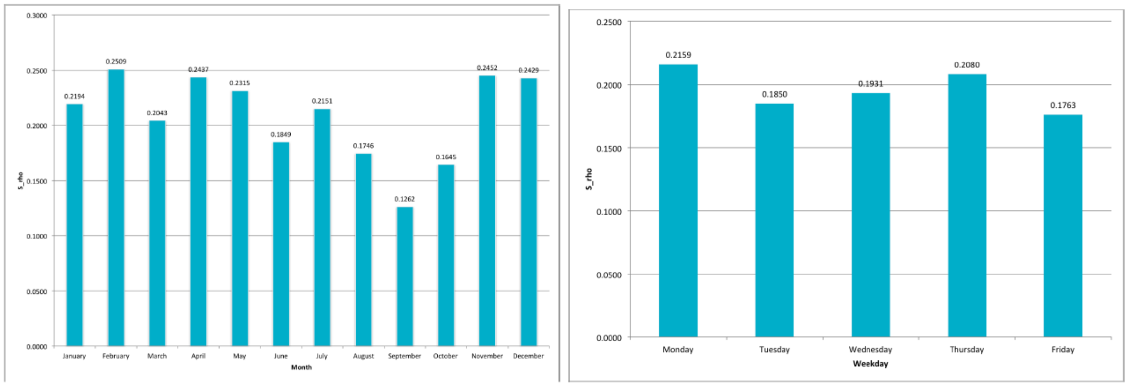


Figure 2.9: Calendar effect of entropy measure

2.4.3 Using Entropy to Test Equality of Densities between Industries

In this part, we apply the same method in the previous section to test the equality of densities for two univariate random variables X and Y, where X and Y are daily returns of Bitcoin and stocks in different industries respectively. The data of 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 201908 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 $t + 1$. 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),

[2]http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data$_{L}ibrary/det_{3}0_{i}nd_{p}ort.html$

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 to test equality of densities for two univariate random variables X and Y by Maasoumi & Racine (2002),

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

where $f_1 = f(x)$ and $f_2 = f(y)$ are the marginal densities of the random variables X and Y, which are daily returns of Bitcoin and stock in different industries respectively.

The result is shown in Table 5 and Figure 9. We can tell from the result that the univariate density of Bitcoin daily return has smallest distance with the univariate density of Coal industry daily return, the S_ρ between these two densities is 0.0039 and statistically significant. The density of Bitcoin daily return also has small distance with densities of Steel, Mines, Games and Txtls industries daily returns, which has S_ρ values of 0.0938, 0.1015, 0.1176 and 0.1176 respectively.

This result makes sense to us due to Bitcoin's energy consumption. O'Dwyer and Malone (2014) and Dilek & Furuncu (2018) indicated that there are some pessimistic argues on Bitcoin and other crytocurrencies because they lack specific center. This could cause a 'balloon-lunacy' financially which would lead to environmental damage due to the energy it consumes. Bitcoin has recently been much on the news be it due to its value or its energy consumption. The rising levels of its energy consumption and the fact that this consumption will continue to increase brings with it a host of negativities. The similarity of nearly 80% of the world's energy consumption on fossil fuels and that this situation is not likely to change in the future brings with it serious problems for the environment. Bitcoin mining spreading in areas where electricity is provided through burning coal causes the already low air quality in these areas to further worsen. The energy consumed due to increased Bitcoin mining is put forward as one of the most important problems impeding Bitcoin's development. The energy consumption of Bitcoin is mainly caused by Bitcoin mining, which always describe the Bitcoin production process. Because in this case mining is used both to confirm processes and to define the people who have put it on record, which is similar with mining for gold. Mining, dependent on the solution of a complicated crypto-puzzle, requires an exorbitant amount of computer power. The main cost of Bitcoin mining is the energy expended to ease the work of accounting done while mining. Internet, hardware maintenance, cables, etc. are all lower than the energy costs (Hayes, 2015). In a case where 400 transactions are done per second, it has been calculated that Bitcoin mining requires 30,582 MW of energy per month (Mishra, 2017). diseases while particles lead to lung diseases. In a study done by Cambridge university, 58% of Bitcoin mining is done in China, followed by the US at 16%. Mining in China, where cheap electricity can be found, involves energy production and consumption based on coal, which negatively effects the environment. Another study done on this topic shows that a Bitcoin center in China continues to depend on coal for the energy consumed by Bitcoin mining (Walt et al., 2017; Hileman & Rauchs, 2017). This situation causes a significant rise in carbon emissions. Bitcoin mining occurring in areas where electricity is obtained from coal means the worsening of air quality. Bitcoin mining is getting even more widespread with every passing day, and it is using up even more energy. The Bitcoin energy consumption theory strengthens our result that Coal, Steel and Mines rank the top 3 in daily return densities' similarity with Bitcoin daily return density. We also plot the densities of Coal, Steel and Mines in Figure 10, Figure 11 and Figure 12. From the figures we can see the shape of densities between Bitcoin and Coal, Steel and Mines are very similar, which indicate that the distance between their densities are small.

Table 2.6: Entropy measure between Bitcoin and different Industries

	Food	Beer	Smoke	Games	Books	Hshld
S_rho	0.2342	0.2294	0.1859	0.1176	0.1478	0.2276
p-value	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16
	Clths	Hlth	Chems	Txtls	Cnstr	Steel
S_rho	0.1372	0.1932	0.1516	0.1176	0.1315	0.0938
p-value	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16
	FabPr	ElcEq	Autos	Carry	Mines	Coal
S_rho	0.1341	0.1415	0.1281	0.1618	0.1015	0.0399
p-value	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16
	Oil	Util	Telcm	Servs	BusEq	Paper
S_rho	0.1287	0.2219	0.2035	0.1715	0.1602	0.1847
p-value	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16
	Trans	Whlsl	Rtail	Meals	Fin	Other
S_rho	0.1595	0.1887	0.1982	0.2039	0.1530	0.1894
p-value	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16	<2.22e-16

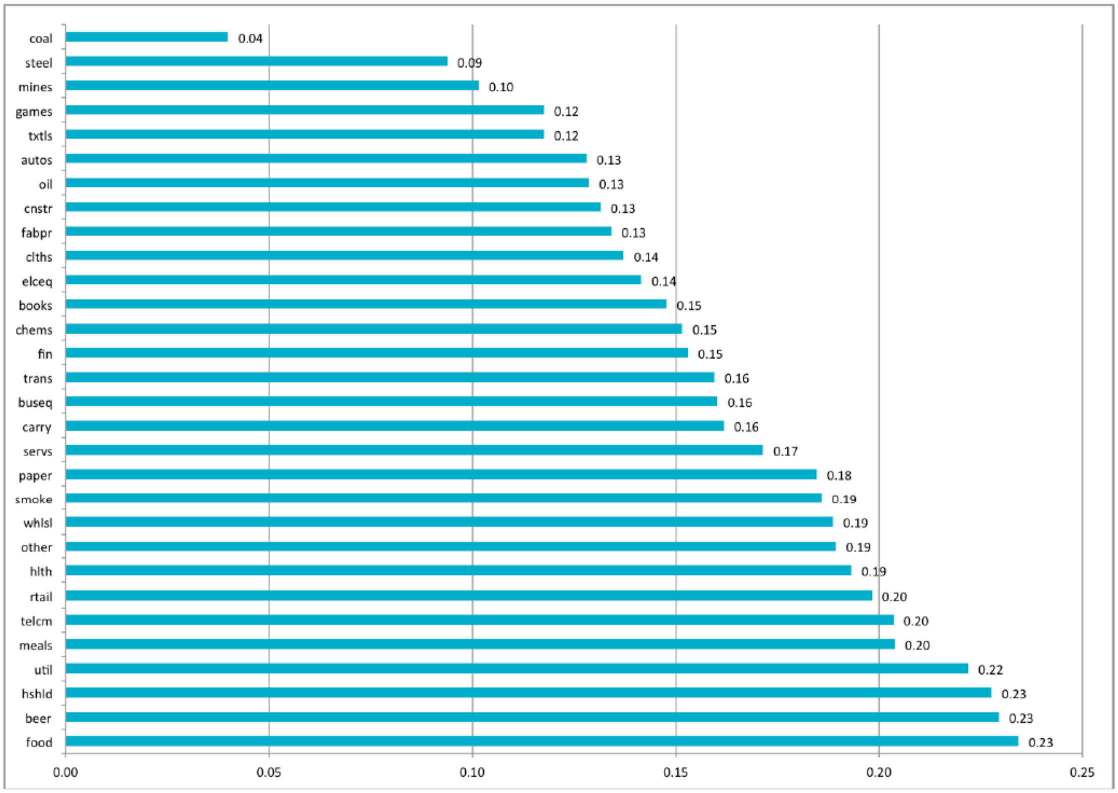


Figure 2.10: Entropy measure between Bitcoin and different Industries

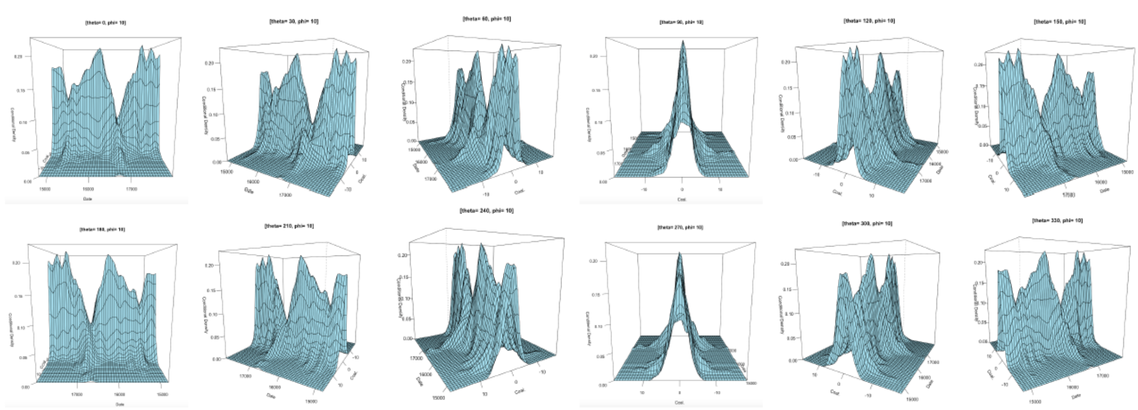


Figure 2.11: Density of Coal

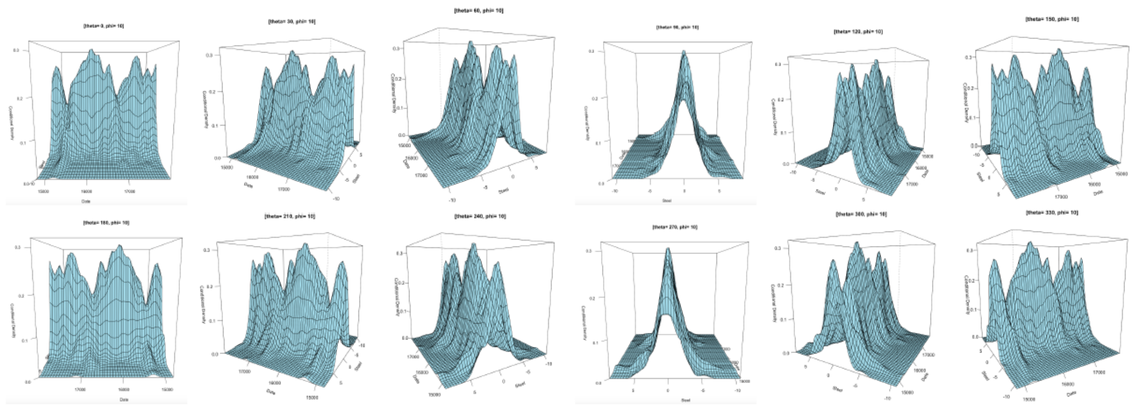


Figure 2.12: Density of Steel

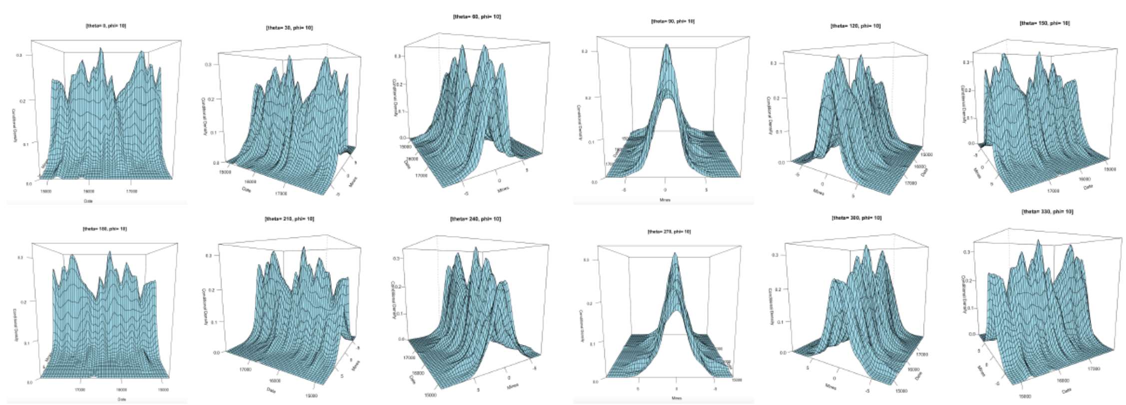


Figure 2.13: Density of Mines

2.4.4 Testing Density Equality Based on Conditional Distribution

In last section, we test the equality of densities for two univariate random variables, return of stock market index and return of cryptocurrency. While in this section, we will turn to conduct nonparametric tests for the equality of conditional density functions. We use w to denote the conditioning discrete variable . w can be a multivariate discrete variable. We use S_w to denote the support of W, and we assume that $P(w) = Pr(W = w)$ is bounded below by a positive constant for all $w \in S_w$. Suppose we have i.i.d data, $\{X_i, U_i\}_{i=1}^{n_1}$, which are random draws from the joint density function $f(x, w)$ along with i.i.d. draws of $\{Y_i, V_i\}_{i=1}^{n_2}$ from the joint density function $g(x, w)$. We use $f(x|w)$ and $g(x|w)$ to denote the conditional density function of x and Y conditional on $U = w$ and $V = w$.

We are testing the null hypothesis: $H_0^c : f(x|w) = g(x|w)$ against the alternative hypothesis: $H_1^c : f(x|w) \neq g(x|w)$.

Define $p_f(w) = Pr(U = w)$ and $p_g(w) = Pr(V = w)$. Note that $p_f(w)$ can differ from $p_g(w)$ here. Using $f(x|w) = f(x, w)/p_f(w)$ and $g(x|w) = g(x, w)/p_g(w)$, we can construct the test statistic based on:

$$J = \sum_{w \in \delta_w} \int [f(x|w) - g(x|w)]^2 dx = \sum_{w \in \delta_w} \int [\frac{f(x,w)^2}{p_f(w)^2} + \frac{g(x,w)^2}{p_g(w)^2} - \frac{2f(x,w)g(x,w)}{p_f(w)p_g(w)}]dx.$$

Let $I_{u_i,w} = I(U_i = w)$ denote an indicator function which equals one if $U_i = w$ and zero otherwise. $I_{v_i,w}$ is similarly defined. We estimate the joint density of $f(x, w)$ and $g(x, w)$ by:

$$\hat{f}(x,w) = \frac{1}{n_1} \sum_{i=1}^{n_1} K_{\gamma,x_i,x} I_{u_i,w}, \qquad \hat{g}(x,w) = \frac{1}{n_2} \sum_{i=1}^{n_2} K_{\gamma,y_i,x} I_{v_i,w}.$$

Also, we estimate $p_f(w)$ and $p_g(w)$ by:

$$\hat{p}_f(w) = \frac{1}{n_1} \sum_{i=1}^{n_1} I(U_i = w), \qquad \hat{p}_g(w) = \frac{1}{n_2} \sum_{i=1}^{n_2} I(V_i = w).$$

Define the leave-one-out empirical functions by:

$$F_{n,-i}(x) = (n_i - 1)^{-1} \sum_{j \neq i}^{n_1} I(X_j \leq x) \qquad G_{n,-i}(x) = (n_2 - 1)^{-1} \sum_{j \neq i}^{n_2} I(Y_j \leq x).$$

Then we obtain the feasible test statistic given by:

$$J_n = \sum_{w\in\delta_w}\int[\frac{\hat{f}(x,w)}{\hat{p}_f^2}dF_{n,-i}(x)+\frac{\hat{g}(x,w)}{\hat{p}_g^2}dG_{n,-i}(x)-\frac{\hat{f}(x,w)}{\hat{p}_f\hat{p}_g}dG_{n,-i}(x)-\frac{\hat{g}(x,w)}{\hat{p}_f\hat{p}_g}dF_{n,-i}(x)]$$

$$=\sum_{w\in\delta_w}\Big\{\sum_i\sum_{j\neq i}[\frac{\bar{K}_{\gamma,x_i,x_j I_{u_i,w} I_{u_j,w}}}{n_1(n_1-1)\hat{p}_f^2}+\frac{\bar{K}_{\gamma,y_i,y_j I_{v_i,w} I_{v_j,w}}}{n_2(n_2-1)\hat{p}_g^2}]$$

$$-\frac{1}{n_1 n_2\hat{p}_f\hat{p}_g}\sum_i\sum_j[\bar{K}_{\gamma,x_i,y_j I_{u_i,w} I_{v_j,w}}+\bar{K}_{\gamma,y_i,x_j I_{v_i,w} I_{u_j,w}}]\Big\}.$$

Li *et al.* (2009) demonstrated that under the null hypothesis,

$$\hat{T}_{n,c} = (n_1 n_2 \hat{h}_1 \hat{h}_2 ... \hat{h}_q)^{1/2}\hat{J}_n/\hat{\sigma}_{n,c} \to N(0,1),$$

where

$$\hat{\sigma}_{n,c}^2 = 2(n_1 n_2 \hat{h}_1...\hat{h}_q)\sum_{w\in\delta_w}\Big[\sum_{i=1}^{n_1}\sum_{j\neq i}^{n_1}\frac{(\bar{K}_{\hat{\gamma},x_i,x_j}I_{u_i,w}I_{u_j,w})^2}{n_1^4\hat{p}_f(w)^4}+\sum_{i=1}^{n_2}\sum_{j\neq i}^{n_2}\frac{(\bar{K}_{\hat{\gamma},y_i,y_j}I_{v_i,w}I_{v_j,w})^2}{n_2^4\hat{p}_g(w)^4}$$

$$+\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\frac{(\bar{K}_{\hat{\gamma},x_i,y_j}I_{u_i,w}I_{v_j,w})^2}{n_1^2 n_2^2\hat{p}_f(w)^2\hat{p}_g(w)^2}+\sum_{i=1}^{n_2}\sum_{j=1}^{n_1}\frac{(\bar{K}_{\hat{\gamma},x_j,y_i}I_{v_i,w}I_{u_j,w})^2}{n_1^2 n_2^2\hat{p}_g(w)^2\hat{p}_f(w)^2}\Big].$$

The test is one-sided rejecting when the test statistic is sufficiently large.

Therefore, we use the test above to analyze the entropy measure of dependence between cryptocurrency returns and stock market indexes returns. In empirical application, we are interested in knowing the distribution of a continuous variable, return, conditional on a discrete variable, asset class. In other words, we are testing the correlation between the conditional distributions of cryptocurrency returns and stock return. Table 7 shows that the value of T_n is smaller when x = Bitcoin and y = NASDAQ, which indicates that the conditional densities of Bitcoin daily returns and NASDAQ daily returns are more similar than other combinations, in other words, the distribution correlation between Bitcoin daily returns and NASDAQ daily returns is the strongest among all combinations. This result is consistent with our result in testing equality of unconditional densities in last section.

Table 2.7: Multivariate entropy density equality test

	T_n	p-value
S&P500 & Bitcoin	224.3750	<2.22e-16
S&P500 & Ethereum	383.8469	<2.22e-16
NASDAQ & Bitcoin	149.7332	<2.22e-16
NASDAQ & Ethereum	313.7069	<2.22e-16

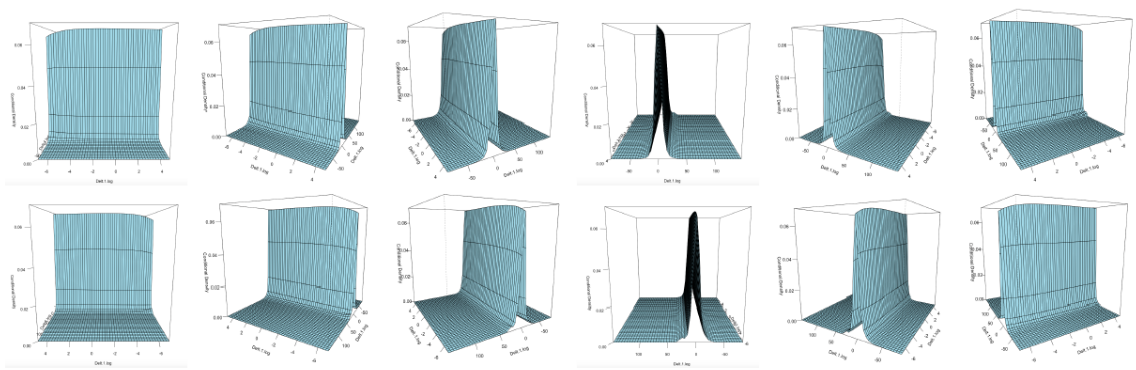


Figure 2.14: Joint Density of S&P500 and Bitcoin

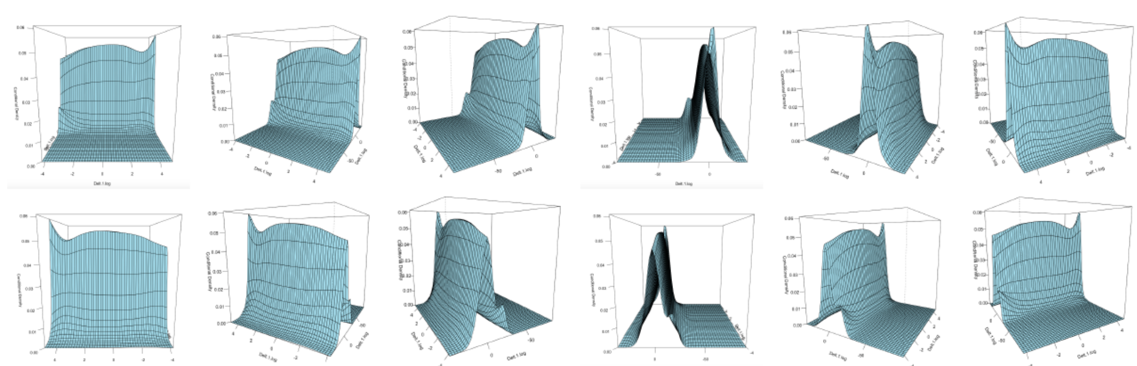


Figure 2.15: Joint Density of S&P500 and Ethereum

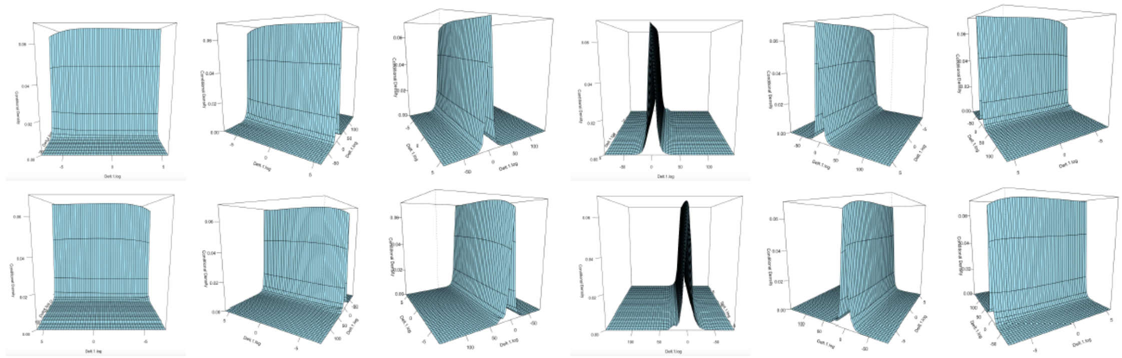


Figure 2.16: Joint Density of Nasdaq and Bitcoin

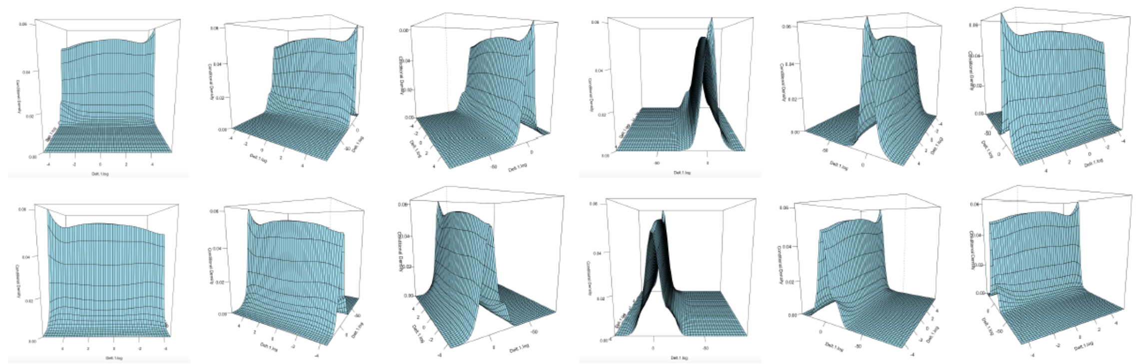


Figure 2.17: Joint Density of Nasdaq and Ethereum

2.4.5 Maximum Entropy Approach to Portfolio Selection

Current Mean-Variance Optimization Approach

Markowitz's (1952) Mean-Variance (MV) optimization is one of the most common formulation of portfolio selection problem. We denote the first two moments of the excess returns $R = (R_1, R_2, ..., R_N)' = (r_1 - r_f, r_2 - f_f, ..., r_N - r_f)'$ on N risky assets as $E(R) = (m_1, m_2, ..., m_N)' = m$, and $Var(R) = ((\sigma_{ij})) = \Sigma$, a $N * N$ matrix, where r_i and r_f denote the return of the i-th, $i = 1, 2, ..., N$ and the risk-free assets, respectively. A portfolio $\pi = (\pi_1, \pi_2, ..., \pi_N)'$ is a vector of weights that represents the investor's relative allocation of the wealth satisfying $\Sigma_{i=1}^N \pi_i = \pi' 1_N$, where 1_N is an $N1$ vector of ones. The Mean-Variance (MV) problem is to choose the portfolio weight vector π to minimize the variance of the portfolio return $Var(\pi' R) = \pi' \Sigma \pi$ subject to a pre-determined target, μ_0 as expected return of the portfolio, i.e.,

$$\underset{\pi}{min}\ \pi' \Sigma \pi \qquad \text{s.t.} \quad E(\pi' R) = \pi' m = \mu_0, \quad \pi' 1_N = 1.$$

Merton (1972) obtained the Lagrange multipliers corresponding to the two constraints in the equation above, respectively, as

$$\gamma = \frac{C\mu_0 - A}{D}, \qquad \nu = \frac{B - A\mu_0}{D},$$

where $A = 1_N' \Sigma^{-1} m$, $B = m' \Sigma^{-1} m$, $C = 1_N' \Sigma^{-1} 1_N$, and $D = BC - A^2$. The solution to the above equation is given by

$$\hat{\pi} = (\frac{\mu_0}{B}) \Sigma^{-1} m$$

at which we have the MV portfolio variance as

$$\sigma_{\hat{\pi}}^2 = \hat{\pi}' \Sigma \hat{\pi} = \frac{C\mu_0^2 - 2A\mu_0 + B}{D}.$$

Therefore, we can write

$$(\frac{D}{C}) \sigma_{\hat{\pi}}^2 - (\mu_0 - \frac{A}{C})^2 = \frac{D}{C^2}.$$

For a given mean and covariance matrix, the MV paradigm provide a very elegant way to achieve an efficient allocation such that higher expected returns can only be achieved by taking on more risk, as it is clear from the efficient frontier equation. Since the MV portfolio $\hat{\pi}$ is derived assuming investor's trade-off between the mean and the variance, the MV portfolio can also be obtained from the following expected utility maximization problem:

$$\underset{\pi}{max}\ E(\pi'R) - \frac{\lambda}{2}Var(\pi'R) \qquad \text{s.t.}\pi'1_N = 1,$$

where λ denotes investor's degree of relative risk aversion.

To construct a portfolio that can best simulate the Bitcoin daily return density, we use the top three industries whose densities have smallest Hellinger distance with the Bitcoin density, which are Steel, Mines and Coal. We first compute the mean, standard deviation and correlation matrix between stock returns and Bitcoin return in Table 8. From the table we can summarize that the volatility of Bitcoin is much higher than the stocks in different industries, in the mean time, the return of Bitcoin is higher then the stocks. This is consistent with the principle of risk-return trade-off that is fundamental to the CAPM. Usually, risk is defined as volatility, which is defined as the standard deviation of the returns. On the other hand, return is either calculated using arithmetic or logarithmic returns.

We construct a portfolio consist of the three industries (Steel, Mines, Coal), our portfolio also has an expected risk-return trade-off. Given weight w_x, w_y and w_z to the three industries respectively, where $w_x \geq 0$, $w_y \geq 0$, $w_z \geq 0$ and $w_x + w_y + w_z = 1$.For the portfolio we get an expected return of

$$\hat{r}_p = w_x\hat{r}_x + w_y\hat{r}_y + w_z\hat{r}_z,$$

and an expected standard deviation of

$$\sigma_p = \sqrt{w_x^2\sigma_x^2 + w_y^2\sigma_y^2 + w_z^2\sigma_z^2 + 2w_xw_y\sigma_{x,y} + 2w_xw_z\sigma_{x,z} + 2w_yw_z\sigma_{y,z}}.$$

We plot the possible portfolio and efficient frontier of the portfolio in Figure 17. We see that the area of possible portfolios has expanded into a third dimension. The colors try to show the two different weights. A yellow color indicates a portfolio consisting mainly of the Steel industry stock, a blue color indicates Mines industry stock, and a red area indicates a portfolio of Coal industry stock. We also see the three single asset portfolios (the three red dots). The efficient frontier is the (upper) edge of all possible portfolios. The efficient frontier can be calculated on its own without the need to simulate thousands of portfolios and then finding the efficient ones. We will look at the case without restrictions (with short-selling allowed). So far we have restricted our portfolios to only contain positive weights (by filtering out negative weights). To calculate the efficient frontier we can use the following closed-form formula that calculates the efficient frontier for a given input of risk (σ) and for some parameters (α, β, γ, δ) using matrix algebra.

$$\hat{r}_{ef}(\sigma) = \frac{\beta}{\alpha} + \sqrt{(\frac{\beta}{\alpha})^2 - \frac{\gamma - \delta\sigma^2}{\alpha}},$$

which is the solution to a quadratic optimization problem. The parameters are given by the following matrix algebra

$$\alpha = 1^T s^{-1} 1,$$

where 1 is a matrix of 1's with a length of the numbers of stocks, s is a matrix of the covariances between the assets (with dimension of $n * n$).

$$\beta = 1^T s^{-1} \overline{ret},$$

where $\overline{ret}$ stands for a vector of average returns for each stock.

$$\gamma = \overline{ret}^T s^{-1} \overline{ret},$$

and lastly δ is given by

$$\delta = \gamma - \beta^2.$$

Given the values for $\alpha = 0.3977$, $\beta = 0.0113$, $\gamma = 0.0015$, and $\delta = 0.0004$ we receive the function for the frontier:

$$\hat{r}_{ef}(\sigma) = \frac{0.0113}{0.3977} \pm \sqrt{(\frac{0.0113}{0.3977})^2 - \frac{0.0015 - 0.0004 * \sigma^2}{0.3977}}.$$

The values for the $+$-part of the function is the upper, efficient frontier, whereas the $-$-part represents the lower, inefficient frontier. The red curve (the upper curve) is the real efficient frontier, whereas the blue curve (the lower curve) stands for an inefficient frontier. This is due to the fact, that we can create a mixture of the three assets that has the same volatility but a higher expected return. As we are able to short-sell (borrow money by selling stocks that we don't own and investing this cash) the efficient frontier doesn't necessarily touch the three assets, nor does it end at the points but extends outwards.

However, portfolios constructed from sample moments of stock returns have proved problematic. The main problems in optimal MV portfolio are that the portfolios are often extremely concentrated on a few asset, which is a contradiction to the notion of diversification, and the out-of-sample performances of the MV portfolios are not very good. It is generally thought that these drawbacks are due to statistical error in estimating the moments that are used as inputs in the MV optimization. These errors are known to change optimal portfolio weights dramatically in such a way that portfolios often involve extreme positions (Jobson and Korkie, 1980).

There have been extensive research on reducing statistical errors in sample mean and covariance matrix. One alternative is the class of shrinkage estimators. Frost and Savarino (1986), Jorion (1986), and Ledoit and Wolf (2003) used shrinkage estimation for the mean and covariance matrix. Shrinkage estimators compensate for the positive (negative) error that tends to be embedded in extremely high (low) estimated coefficients by pulling them downward (upward) and prevent extreme positions in portfolio selection. Since shrinkage estimators are based on the empirical Bayesian approaches, a particular prior distribution should be assumed to derive those estimators. Although some prior distributions used in the empirical Bayes estimation are known to work well, there is no systematic way to choose a prior distribution. For example, Jorion (1986) used an informative conjugate prior and derived the multivariate normal predictive distribution with the mean of minimum variance portfolio as the target mean. Frost and Savarino (1986) adapted a normal-wishart conjugate prior and derived multivariate Student's t predictive density. In their simulation study, they assumed that means, variances and correlations for all the assets are the same, so that their target mean and covariance matrix are those of equally weighted portfolio. As a result, it is very hard to achieve a certain shrinkage target preferred by asset managers, for example, a capitalization weighted portfolio.

Therefore, we propose a method that ensures shrinkage towards maximum diversification of portfolio weights using a information theoretic approach.

Table 2.8: Sample means, variances, and correlation matrix

	Steel	Mines	Coal	Bitcoin
Mean	0.0284	0.0157	-0.0431	0.3791
Variance	1.7373	1.6703	2.8214	7.1256
Correlation	1.0000			
	2.2574	1.0000		
	3.1939	2.9771	1.0000	
	0.5610	0.5667	0.8300	1.0000

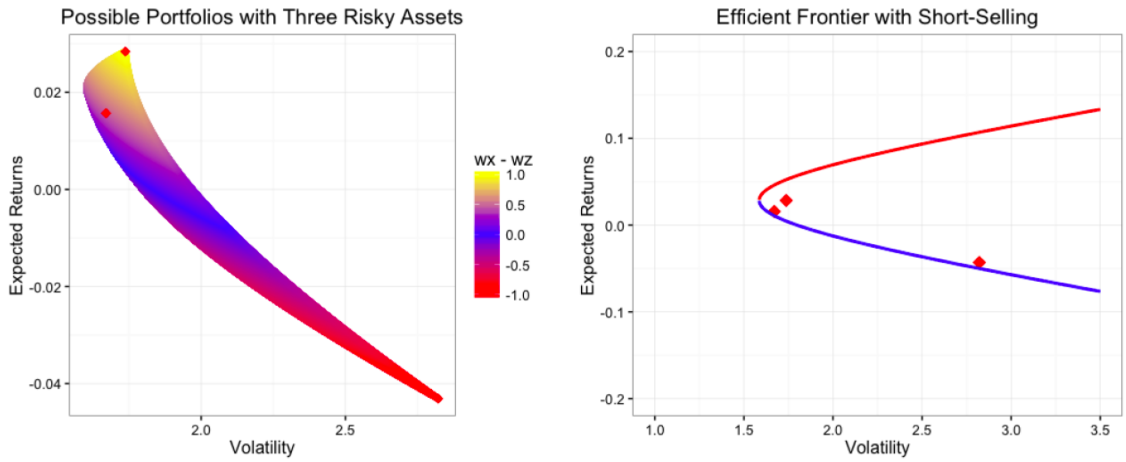


Figure 2.18: Possible portfolios and efficient frontier

Information Theoretic Approach

In this part, we propose a maximum entropy approach to portfolio selection. It can go beyond the quadratic utility functions and ensures shrinkage towards maximum diversification of portfolio weights. Our objective function, the Kullback–Leibler information criteria is defined as pseudo distance between two probability distributions (portfolio weights), $p = (p_1, p_2, ..., p_N)'$ and $q = (q_1, q_2, ..., q_N)'$:

$$KLIC(p, q) = \sum_{i=1}^{N} p_i ln(p_i/q_i).$$

The KLIC is also known as the cross-entropy (CE) measure (Golan et al., 1996). If one minimizes the CE measure with q as the reference distribution that satisfies certain constraints, one can get a solution, $\hat{p}$, closest to q. If we set $q = (1/N, 1/N, ..., 1/N)'$, uniform distribution, then $KLIC(p, q)$ is same as negative Shannon's (1948) entropy measure. Since maximizing Shannon's entropy subject to some moment constraints implies estimating p that is the closest to the uniform distribution (i.e., equally weighted portfolio), well-diversified optimal portfolio can be achieved.

In order to incorporate problems of imprecision of sample moments estimates, we define the confidence interval of maximized expected utility values which lead to inequality constraints to our optimization procedures. This confidence interval can be interpreted as the degree of uncertainty for the sample moments estimates, and can be estimated by resampling methods such as bootstrap or Monte Carlo approaches.

First, we will introduce the entropy measures. A discrete probability distribution $p = (p_1, p_2, ..., p_N)'$ of a random variable taking N values provides a measure of uncertainty (disorder) regarding that random variable. In the information theory literature, this measure of disorder is called entropy. Entropy measures have been extensively used in econometrics, and for more on this see Maasoumi (1993), Golan et al. (1996), Ullah (1996), and Bera and Bilias (2002).

A portfolio allocation $\pi = (\pi_1, \pi_2, ..., \pi_N)'$ among N risky assets, with properties $\pi_i \geq 0, \quad i = 1, 2, ..., N$ and $\sum_{i=1}^{N} \pi_i = 1$, has the structure of a proper probability distribution. We will use the Shannon entropy (SE) measure

$$SE(\pi) = -\sum_{i=1}^{N} \pi_i ln \pi_i$$

as a measure of portfolio diversification. When $\pi_i = 1/N$ for all i, $SE(\pi)$ has its maximum value lnN. The other extreme case occurs when $\pi_i = 1$ for one i, and $= 0$ for the rest, then $SE(\pi) = 0$. Therefore, SE that provides a good measure of disorder in a system or expected information in a probability distribution, can be taken as a measure of portfolio diversification. In financial applications, portfolios are generally evaluated in terms of their degree of diversification using the SE measure after portfolios are obtained using different selection procedures (Fernholz, 2002; Hoskisson et al., 1993; Lubatkin et al., 1993). We put the entropy itself in the objective function so as to obtain maximum diversity in a portfolio allocation. It is clear that when we maximize $SE(\pi)$ we shrink the portfolio towards an equally weighted portfolio, namely, $N^{-1}1 = (1/N, 1/N, ..., 1/N)'$. We wil also consider a more general objective function. Suppose a portfolio weight changes from π_i to q_i, then the change in entropy is $-lnq_i - (-ln\pi_i) = ln(\pi_i/q_i)$. Taking averge of $ln(\pi_i/q_i)$ with π_i's as weights we end up with the notion of CE, $CE(\pi, q = KLIC(\pi, q)$. It is clear that when $q = (1/N, 1/N, ..., 1/N)'$, $CE(\pi, q) = \sum_{i=1}^{N} \pi_i ln\pi_i - lnN$. Therefore, maximization of SE is a special case of CE minimization with respect to an equally weighted portfolio. In our analysis, we will emphasize the minimization of $CE(\pi, q)$ for a given q as a reasonable opportunity set for an investor. Thus, starting from an initial portfolio allocation q, through minimization of CE we can obtain a more diversified portfolio. Golan et al. (1996) showed that

$$CE(\pi, q) = \sum_{i=1}^{N} \pi_i ln(\pi_i/q_i) \approx \sum_{i=1}^{N} \frac{1}{q_i}(\pi_i - q_i)^2 \qquad \text{for} \quad q_i > 0.$$

Thus, we adjust small allocations of the initial portfolio q more than the large ones, possibly resulting in a more diversified portfolio.

To incorporate estimation imprecision of the mean and covariance (as in Bayes-Stein estimation), we need more general constraint. In general, we consider the following minimization problem:

$$\min_{\pi} CE(\pi|q) = \sum_{i=1}^{N} \pi_i ln(\pi_i/q_i)$$

subject to

$$EU(\pi, R, \lambda) \geq \tau, \qquad \pi \geq 0, \qquad \text{and} \quad \pi' 1_N = 1,$$

where $U(\pi, R, \lambda)$ is an utility function, λ is the risk aversion parameter, and τ reflects investor's strength of belief in the estimated expected utility values, which we elaborate further below. We assume that $N * 1$ random vector R has a distribution function $F(R)$ with density $f(R)$. To see the significance of τ, we define

$$\xi \equiv EU(\breve{\pi}, R, \lambda),$$

where $\breve{\pi} = (\breve{\pi}_1, \breve{\pi}_2, ..., \breve{\pi}_N)$ satisfies following expected utility maximization,

$$\breve{\pi} = \text{argmax } EU(\pi, \tilde{R}, \lambda)$$

subject to

$$\pi'1 = 1, \qquad \text{and} \quad \pi \geq 0,$$

where $\tilde{R}$ is a random sample of size T drawn from the empirical distribution $\hat{F}(R)$. As we mentioned previously, estimation imprecision of the sample moments can be measured directly by resampling methods. Solving the optimization problem using B sets of samples leads to B portfolios, $\breve{\pi}_b, \quad b = 1, 2, ..., B$. The investor's strength of belief parameter τ can also be related to the degree of shrinkage and be expressed as, as say the r-th quantile of the distribution of $\xi, \quad 0 < r < 1$, i.e.,

$$\tau = G^{-1}(r) \equiv \xi_r,$$

where $G()$ is distribution of function of ξ. Thus, the first inequality constraint, $EU(\pi, R, \lambda) \geq \tau$, can be represented as a confidence interval. $I = [\xi_r, \xi^U]$, where ξ^U is the same as the maximized expected utility of MV efficient portfolio given λ if $EU()$ is the quadratic expected utility function. This is due to the fact that when there is no estimation error, the maximized expected utility evaluated at these exact moments dominates all values generated by $\breve{\pi}_b$, $b = 1, 2, ..., B$.

The confidence interval has a nice interpretation as a measure of uncertainty (Bewley, 1988). Suppose an investor has high uncertainty aversion in the portfolio selection problem. Then, he/she will select relatively low τ, i.e., ξ_r with a small value of r, and use a $(1-r)\%$ confidence interval. Since $\tau \equiv \xi_r$ represents an investor's strength of belief, we can correspond ξ_r with a large value of r, with investor who has less uncertainty in estimation, and vice-versa. Garlappi et al. (2004) used the notion of the confidence interval to explain investor's aversion toward uncertainty using a multi-prior approach, and showed that their estimated portfolio weights shrink toward the weights of minimum variance portfolio more than those of empirical Bayes-Stein portfolio. While recent studies based on the empirical Bayes-Stein estimator tried to estimate admissible moments at the first stage and then optimize the portfolio weights by the MV principle, weights achieved by minimizing CE objective function subject to sets of constraints are shrunk directly to an appropriate prior weights, q. Moreover, as Frost and Savarino (1986) emphasized, there is no certain way to select a particular informative prior in Bayesian decision rules. One can readily choose alternative informative priors for the Bayes-Stein estimator and obtain different type of shrinkage estimators for portfolio weights by calculating somewhat complex predictive density. However, instead of choosing alternative informative priors, one can choose an appropriate prior weight vector q, and minimize the CE measure to estimate portfolio weights which also has the shrinkage interpretation. Thus, we can say that CE measure works directly as shrinkage estimator of portfolio weights in asset allocation problem.

We can use bootstrap or Monte Carlo methods to estimate a distribution of ξ, i.e., resampling TxN samples for B times from the empirical distribution, $\hat{F}(F)$. Let these resampled series be $\tilde{R}_{(b)}, \quad b = 1, 2, ..., B$. Then, $\breve{\pi}_{(b)}$ and $\xi_{(b)}$ can be calculated as follows,

$$\breve{\pi}_{(b)} = \text{argmax}[\pi'\tilde{m}_{(b)} - \frac{\lambda}{2}\pi'\tilde{\Sigma}_{(b)}\pi],$$

$$\xi_{(b)} = \breve{\pi}_{(b)}\hat{m} - \frac{\lambda}{2}\breve{\pi}'_{(b)}\hat{\Sigma}\breve{\pi}_{(b)},$$

where $\hat{m}$ and $\hat{\Sigma}$ are the sample mean and sample covariance matrix estimated from original return data R, and $\tilde{m}_{(b)}$ and $\tilde{\Sigma}_{(b)}$ are calculated from simulated data $\tilde{R}_{(b)}$. The empirical distribution of ξ can be estimated based on $\xi_{(b)}, \quad b = 1, 2, ..., B$. Then, the CE minimization problem can be written as

$$\min \sum_{i=1}^{N} \pi_i ln(\pi_i / q_i)$$

subject to

$$\pi' \hat{m} - \frac{\lambda}{2} \pi' \hat{\Sigma} \pi \geq \hat{G}^{-1}(r), \qquad \pi \geq 0, \quad \text{and} \quad \pi' 1_N = 1,$$

where $\hat{G}()$ denotes the empirical distribution function of ξ. Under the assumption of smooth expected utility function, it is straightforward to solve the optimization problem minimize by classical gradient based routine.

Figure 18 shows contour curves of $-\sum_{i=1}^{N} \pi_i ln(\pi_i)$ and $-\sum_{i=1}^{N} \pi_i ln(\pi_i/\pi_i^{min})$ respectively. The left figure shows the contour curves of $-\sum_{i=1}^{N} \pi_i ln(\pi_i)$ for the three assets ($N = 3$) on the daily portfolio standard deviation-mean plane. We consider every possible combination of weights π_1, π_2 and π_3, each taking 50 equally spaced values in (0, 1), and satisfying $\sum_{i=1}^{N} \pi_i = 1$. The upper envelope curve in Figure 18 (left) corresponds to the set of MV efficient portfolio. The point where $\sum_{i=1}^{N} \pi_i ln(\pi_i)$ takes the highest value represents equally weighted portfolio. The smoothness of each contour curve ensures existence of a unique solution if we are to solve the minimization problem with $q = N^{-1}1$. Figure 18 (right) shows contour curves of $-\sum_{i=1}^{N} \pi_i ln(\pi_i/\pi_i^{min})$, where π^{min} is portfolio weights for minimum variance portfolio. We can see that the largest value of the function corresponds to minimum variance portfolio. Since minimum variance portfolio does not take account of the portfolio mean value, contour graph on the right hand side is sensitive to the mean values compared to that on the left hand side. Thus, by minimizing CE with $q = \pi^{min}$, it shrinks toward minimum variance portfolio and at the same time takes care of the portfolio mean values.

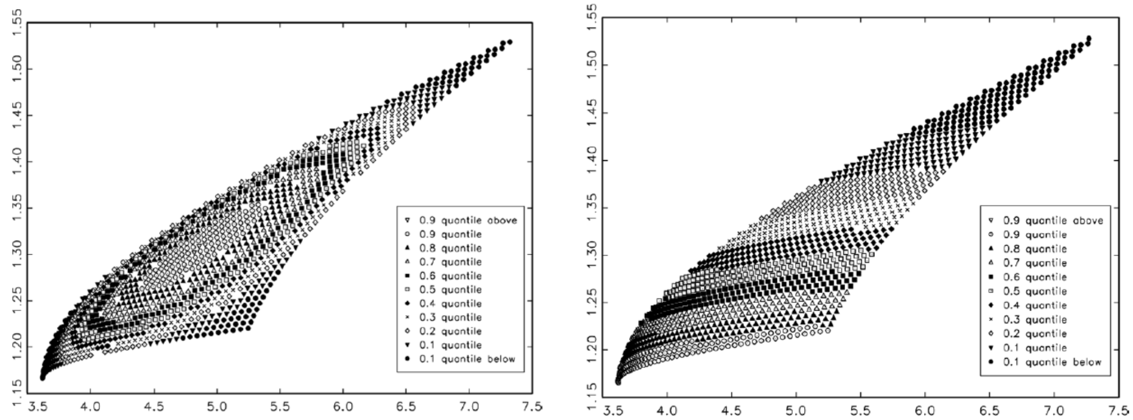


Figure 2.19: Contour curves

There are several advantages in our information theoretic approach: (I) In our approach, our utility function can be any form instead of limited to quadratic utility functions. (II) While previous articles primarily dealt with shrinkage estimators for the mean and covariance matrix to obtain more well-behaved optimal portfolios, we directly shrink portfolio composition (p) towards predetermined target portfolio weights (q) that are of interest to asset managers; (III) Most asset managers are not allowed to sell short (i.e. the portfolio weights cannot be negative) in the real world. Since constructed portfolio weights obtain through the maximum entropy (ME) approach are in the form of "probabilities", the weights are certainly nonnegative. However, negative portfolio weights, when they are appropriate, for example, in case of hedge funds, can also be obtained using the generalized cross entropy (GCE) framework; (IV) Since the mean and covariance matrix should be estimated, one usually has only partial information. It is known that if sample sizes of individual returns are not large enough compare to the number of stocks, sample covariance matrix tends to be very imprecise. By minimizing the CE measure subject to certain well defined constraints, one can extract useful information from the sample mean and covariance matrix.

2.5 Conclusion

This chapter investigates the similarity structure between cryptocurrency and stock returns. We first provide the statistical properties of our proposed models and study in detail their forecasting performance and adequacy by time series models (ARIMA model and GARCH model). The entropy profiles method and time series models play different roles in forecasting cryptocurrency returns volatility. These two methods are not competing with each other, but complimentary. The time series method's advantage is that it can give us a exact form of models, while the entropy profiles method is a nonparametric approach which can get rid of misspecification problem. Then, we calculate nonparametric entropy metrics to test the density equality between two cryptocurrencies (Bitcoin and Ethereum) and two stock market indexes (S&P500 and NASDAQ). We find density similarity between Bitcoin and NASDAQ stock market index. Next, we analyze the fluctuation of entropy measure within months and weekdays, and find that the distance is smaller in September and stable among weekdays. Then we calculate the entropy metrics between Bitcoin and different industries of stocks, and find that the top three industries which have similar densities with Bitcoin is Coal, Steel and Mines. Therefore, we can construct a portfolio using these three industries of stocks, which can best simulate the performance of Bitcoin return. Last, we discuss the current mean-variance portfolio optimization method, then propose a maximum entropy approach to portfolio selection, which ensures shrinkage towards maximum diversification of portfolio weights. Our findings will be useful in exploring the prediction of cryptocurrency returns based on stock market performance.

Our main conclusion is that indeed cryptocurrency represents an asset class that can be assessed using simple finance tools. At the same time, cryptocurrency comprise an asset class which is radically different from traditional asset classes. The statistics S_ρ and maximum entropy approach are useful in forecasting returns of cryptocurrency based on stock market performance.

Bibliography

[1] Al-Khazali, Osamah, Bouri Elie, and David Roubaud. "The impact of positive and negative macroeconomic news surprises: Gold versus Bitcoin." Economics Bulletin 38, no. 1 (2018): 373-382.

[2] Balcilar, Mehmet, Elie Bouri, Rangan Gupta, and David Roubaud. "Can volume predict Bitcoin returns and volatility? A quantiles-based approach." Economic Modelling 64 (2017): 74-81.

[3] Balcombe, Kelvin, and Iain Fraser. "Do bubbles have an explosive signature in markov switching models?." Economic Modelling 66 (2017): 81-100.

[4] Bariviera, Aurelio F. "The inefficiency of Bitcoin revisited: A dynamic approach." Economics Letters 161 (2017): 1-4.

[5] Bariviera, Aurelio F., María José Basgall, Waldo Hasperué, and Marcelo Naiouf. "Some stylized facts of the Bitcoin market." Physica A: Statistical Mechanics and its Applications484 (2017): 82-90.

[6] Baur, Dirk G., Thomas Dimpfl, and Konstantin Kuck. "Bitcoin, gold and the US dollar–A replication and extension." Finance research letters 25 (2018): 103-110.

[7] Begušić, Stjepan, Zvonko Kostanjčar, H. Eugene Stanley, and Boris Podobnik. "Scaling properties of extreme price fluctuations in Bitcoin markets." Physica A: Statistical Mechanics and its Applications 510 (2018): 400-406.

[8] Bera, Anil K., and Yannis Bilias. "The MM, ME, ML, EL, EF and GMM approaches to estimation: a synthesis." Journal of Econometrics 107, no. 1-2 (2002): 51-86.

[9] Berrett, Thomas B., Richard J. Samworth, and Ming Yuan. "Efficient multivariate entropy estimation via K-nearest neighbour distances." The Annals of Statistics 47, no. 1 (2019): 288-318.

[10] Bewley, Truman F. "Knightian decision theory and econometric inference." (1988).

[11] Bouri, Elie, Georges Azzi, and Anne Haubo Dyhrberg. "On the return-volatility relationship in the Bitcoin market around the price crash of 2013." Economics 11, no. 1 (2017).

[12] Bouri, Elie, Naji Jalkh, Peter Molnár, and David Roubaud. "Bitcoin for energy commodities before and after the December 2013 crash: diversifier, hedge or safe haven?." Applied Economics 49, no. 50 (2017): 5063-5073.

[13] Bouri, Elie, Peter Molnár, Georges Azzi, David Roubaud, and Lars Ivar Hagfors. "On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier?." Finance Research Letters 20 (2017): 192-198.

[14] Bouri, Elie, Rangan Gupta, Aviral Kumar Tiwari, and David Roubaud. "Does Bitcoin hedge global uncertainty? Evidence from wavelet-based quantile-in-quantile regressions." Finance Research Letters 23 (2017): 87-95.

[15] Bouri, Elie, Syed Jawad Hussain Shahzad, and David Roubaud. "Co-explosivity in the cryptocurrency market." Finance Research Letters 29 (2019): 178-183.

[16] Catania, Leopoldo, and Stefano Grassi. "Modelling crypto-currencies financial time-series." Available at SSRN 3028486(2017).

[17] Chan, Stephen, Jeffrey Chu, Saralees Nadarajah, and Joerg Osterrieder. "A statistical analysis of cryptocurrencies." Journal of Risk and Financial Management 10, no. 2 (2017): 12.

[18] Chu, Jeffrey, Stephen Chan, Saralees Nadarajah, and Joerg Osterrieder. "GARCH modelling of cryptocurrencies." Journal of Risk and Financial Management 10, no. 4 (2017): 17.

[19] Conrad, Christian, Anessa Custovic, and Eric Ghysels. "Long-and short-term cryptocurrency volatility components: A GARCH-MIDAS analysis." Journal of Risk and Financial Management 11, no. 2 (2018): 23.

[20] Dickey, David A., and Wayne A. Fuller. "Distribution of the estimators for autoregressive time series with a unit root." Journal of the American statistical association 74, no. 366a (1979): 427-431.

[21] DİLEK, Şerif, and Yunus Furuncu. "Bitcoin mining and its environmental effects." Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi 33, no. 1 (2019): 91-106.

[22] Dyhrberg, Anne Haubo. "Bitcoin, gold and the dollar–A GARCH volatility analysis." Finance Research Letters 16 (2016): 85-92.

[23] Fernholz, E. Robert. "Stochastic portfolio theory." In Stochastic portfolio theory, pp. 1-24. Springer, New York, NY, 2002.

[24] Frost, Peter A., and James E. Savarino. "An empirical Bayes approach to efficient portfolio selection." Journal of Financial and Quantitative Analysis 21, no. 3 (1986): 293-305.

[25] Frost, Peter A., and James E. Savarino. "An empirical Bayes approach to effi-

cient portfolio selection." Journal of Financial and Quantitative Analysis 21, no. 3 (1986): 293-305.

[26] Garlappi, Lorenzo. "Risk premia and preemption in RD ventures." Journal of Financial and Quantitative Analysis 39, no. 4 (2004): 843-872.

[27] Golan, Amos, George G. Judge, and Douglas Miller. Maximum entropy econometrics. No. 1488. Iowa State University, Department of Economics, 1996.

[28] Granger, C. W., and Esfandiar Maasoumi. "A dependence metric for nonlinear time series." In Econometric Society World Congress 2000 Contributed Papers, no. 0421. Econometric Society, 2000.

[29] Granger, Clive W., Esfandiar Maasoumi, and Jeffrey Racine. "A dependence metric for possibly nonlinear processes." Journal of Time Series Analysis 25, no. 5 (2004): 649-669.

[30] Granger, Clive WJ, Bwo-Nung Huangb, and Chin-Wei Yang. "A bivariate causality between stock prices and exchange rates: evidence from recent Asianflu" The Quarterly Review of Economics and Finance 40, no. 3 (2000): 337-354.

[31] Gulko, Les. "The entropy theory of stock option pricing." International Journal of Theoretical and Applied Finance 2, no. 03 (1999): 331-355.

[32] Hayes, Adam. "The decision to produce altcoins: Miners' arbitrage in cryptocurrency markets." Available at SSRN 2579448 (2015).

[33] Hileman, Garrick, and Michel Rauchs. "Global cryptocurrency benchmarking study." Cambridge Centre for Alternative Finance 33 (2017): 33-113.

[34] Hoskisson, Robert E., Michael A. Hitt, Richard A. Johnson, and Douglas D. Moesel. "Construct validity of an objective (entropy) categorical measure of diversification strategy." Strategic management journal 14, no. 3 (1993): 215-235.

[35] Jaynes, Edwin T. ”On the rationale of maximum-entropy methods.” Proceedings of the IEEE 70, no. 9 (1982): 939-952.

[36] Jobson, J. David, and Bob Korkie. ”Estimation for Markowitz efficient portfolios.” Journal of the American Statistical Association 75, no. 371 (1980): 544-554.

[37] Jorion, Philippe. ”Bayes-Stein estimation for portfolio analysis.” Journal of Financial and Quantitative analysis 21, no. 3 (1986): 279-292.

[38] Katsiampa, Paraskevi. ”Volatility estimation for Bitcoin: A comparison of GARCH models.” Economics Letters 158 (2017): 3-6.

[39] Kullback, Solomon, and Richard A. Leibler. ”On information and sufficiency.” The annals of mathematical statistics 22, no. 1 (1951): 79-86.

[40] Lahmiri, Salim, and Stelios Bekiros. ”Chaos, randomness and multi-fractality in Bitcoin market.” Chaos, solitons fractals106 (2018): 28-34.

[41] Lahmiri, Salim, Stelios Bekiros, and Antonio Salvi. ”Long-range memory, distributional variation and randomness of bitcoin volatility.” Chaos, Solitons Fractals 107 (2018): 43-48.

[42] Ledoit, Olivier, and Michael Wolf. ”Improved estimation of the covariance matrix of stock returns with an application to portfolio selection.” Journal of empirical finance 10, no. 5 (2003): 603-621.

[43] Lubatkin, Michael, Hemant Merchant, and Narasimhan Srinivasan. ”Construct validity of some unweighted product-count diversification measures.” Strategic Management Journal 14, no. 6 (1993): 433-449.

[44] Maasoumi 1, Esfandiar. ”A compendium to information theory in economics and econometrics.” Econometric reviews 12, no. 2 (1993): 137-181.

[45] Maasoumi, Esfandiar, and Jeff Racine. "Entropy and predictability of stock market returns." Journal of Econometrics107, no. 1-2 (2002): 291-312.

[46] Markowitz, Harry. "The utility of wealth." Journal of political Economy 60, no. 2 (1952): 151-158.

[47] Mishra, Sailendra Prasanna, Varghese Jacob, and Suresh Radhakrishnan. "Energy consumption–Bitcoin's Achilles heel." Available at SSRN 3076734 (2017).

[48] O'Dwyer, Karl J., and David Malone. "Bitcoin mining and its energy footprint." (2014): 280-285.

[49] Osterrieder, Joerg, and Julian Lorenz. "A statistical risk assessment of Bitcoin and its extreme tail behavior." Annals of Financial Economics 12, no. 01 (2017): 1750003.

[50] Pichl, Lukáš, and Taisei Kaizoji. "Volatility analysis of bitcoin." Quantitative Finance and Economics 1, no. 4 (2017): 474-485.

[51] Rényi, Alfréd. "On measures of entropy and information." In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics, vol. 4, pp. 547-562. University of California Press, 1961.

[52] Renyi, Alfred. "On the number of endpoints of a k-tree." Studia Sci. Math. Hungar 5 (1970): 5-10.

[53] Robinson, Peter M. "Consistent nonparametric entropy-based testing." The Review of Economic Studies 58, no. 3 (1991): 437-453.

[54] Segnon, Mawuli, and Stelios Bekiros. "Forecasting Volatility in Cryptocurrency Markets." Center for Quantitative Econmics79, no. 1 (2019): 1-37.

[55] Shannon, Claude Elwood. "A mathematical theory of communication." The Bell system technical journal 27, no. 3 (1948): 379-423.

[56] Silverman, Bernard W. Density estimation for statistics and data analysis. Routledge, 2018.

[57] Skaug, Hans Julius, and Dag Tjøstheim. "Testing for serial independence using measures of distance between densities." In Athens Conference on Applied Probability and Time Series Analysis, pp. 363-377. Springer, New York, NY, 1996.

[58] Stavroyiannis, Stavros. "Value-at-risk and related measures for the Bitcoin." The Journal of Risk Finance (2018).

[59] Thies, Sven, and Peter Molnár. "Bayesian change point analysis of Bitcoin returns." Finance Research Letters 27 (2018): 223-227.

[60] Ullah, Aman. "Entropy, divergence and distance measures with econometric applications." Journal of Statistical Planning and Inference 49, no. 1 (1996): 137-162.

[61] Walt E. V. Gao G. Biesheuvel T, and Warren H. (2017) Coal Is Fueling Bitcoin's Meteoric Rise, Bloomberg, 15 December 2017.

[62] Zheng, John Xu. "A consistent test of conditional parametric distributions." Econometric Theory 16, no. 5 (2000): 667-691.

Chapter 3

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

In this chapter, 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.[1]

Keywords:

Cryptocurrency, Bitcoin, Entropy, Co-dependence, COVID-19, Vaccine

[1]This essay has been published: Maasoumi, Esfandiar, and Xi Wu. "Contrasting Cryptocurrencies with Other Assets: Full Distributions and the COVID Impact." Journal of Risk and Financial Management 14, no. 9 (2021): 440.

3.1 Introduction

Since the emergence of Bitcoin based on blockchain technology in 2018, global financial markets have witnessed the birth and rapid rise of cryptocurrencies as a new asset class. Cryptocurrencies are based on fundamentally new technologies, the potential of which highly anticipated but not fully understood. In their current form, however, cryptocurrencies 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), Baur 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.

Since the global financial crisis in 2008, the financial markets seems to have stabilized until the World Health Organization (WHO) declared COVID-19 as a pandemic in 2020. The global economy once again fell into panic: Consumer consumption has changed, and the types and quantities of products have also changed. Many companies are trying to escape from the panic by the changes in business and marketing, but they are still unable to recover their business or reduce their financial burden easily.

In addition, the unemployment rate and the COVID-19 incidence rate are still not improving. As the liquidity of capital decreases in many ways due to these changes, there is also a phenomenon of seeking new investment destinations. Recent studies show that the effect of the recently developed COVID-19 vaccine will improve real assets and financial assets (David et al., 2021; Del Giudice et al., 2020; Gherghina et al., 2020). The cryptocurrency market also shows a sharp decline due to COVID-19 after booming in 2017. Therefore, from an investment perspective, we aim to examine the evolution of the complexity of the cryptocurrency market and analyze the characteristics of the market from the past bull market to the present COVID-19 pandemic outbreak to help policymakers and decision-makers ensure future stability.

Recent studies also examined the performance of cryptocurrencies under COVID-19 pandemic. Vukovic et al. (2021) developed a unique COVID-19 global composite index that measures COVID-19 pandemic time-variant movements on each day. Sarkodie et al. (2021) investigated the implication of COVID-19 outcomes on market prices of several leading cryptocurrencies. Naeem et al. (2021) quantified the spillover effects among seven cryptocurrencies to explore the spillover characteristics cryptocurrencies, and discovered that Bitcoin, Litecoin, and Ripple are the dominant transmitters to return spillover. Kim and Lee (2021) investigated the evolution of the complexity of the cryptocurrency market and analyze the characteristics from the past bull market in 2017 to the present the COVID-19 pandemic, and concluded that financial market unpredictability is increasing by the ongoing health crisis. These studies inspired us on investigating deeper on how the density similarities between cryptocurrencies, stocks and industry groups will be affected by COVID-19 outbreak.

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 chapter 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 cryptocurrency returns, in this chapter we use entropy profiles of different asset classes and indices, as well as the cryptocurrencies. 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 cryptocurrency which have large volume and relatively long histories. We use nonparametric entropy metrics to test equality between cryptocurrency 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 chapter we examine the general definition of dependence between cryptocurrency return and stock market returns. Stochastic independence is tested and degree of dependence is measured with entropy metrics.

The rest of the chapter 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 cryptocurrencies (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 cryptocurrencies. Section 5 extended our previous analysis into longer period, and conducted a three-period analysis to study the effect of vaccine. Section 6 provides the concluding remarks.

3.2 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[2]. 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 cryptocurrency 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 cryptocurrency market data and stock market indexes, we calculate the daily log return using adjusted close price:

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

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 cryptocurrency 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.

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

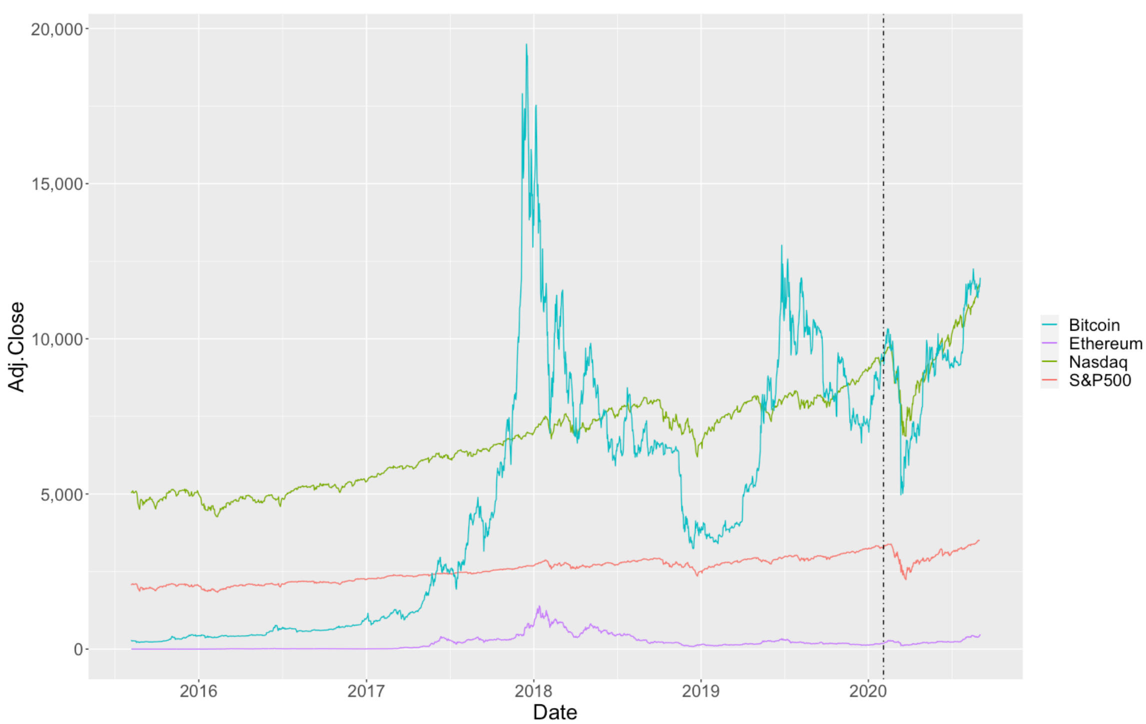
We applied the Augmented-Dicker-Fuller (ADF) unit-root test, which suggests stationarity of the log-returns. An ADF test tests the null hypothesis that a unit root is present in a time series sample. The alternative hypothesis is different depending on which version of the test is used, but is usually 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 cryptocurrency 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 S&P 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 cryptocurrencies 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 3.1: 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.*



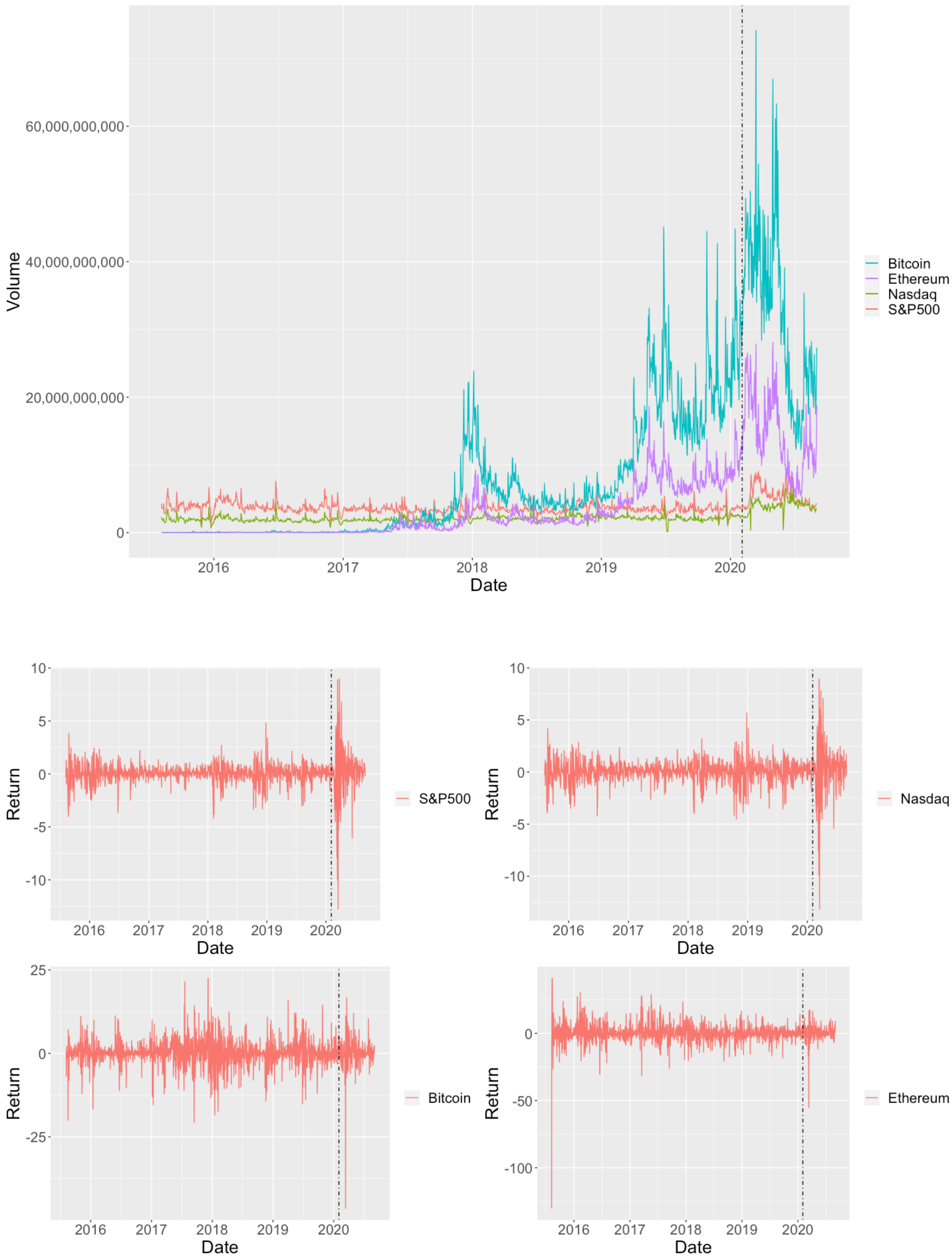


Figure 3.1: Plot of price, volume and daily log-returns

3.3 Entropy Profiles Method

3.3.1 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:

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

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

$$H(x) = -\int p(x)log(p(x))dx,$$

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

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

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

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

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

$$H_\alpha^R(p) = \frac{1}{1-\alpha} log \sum_k p_k^\alpha.$$

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

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

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).

3.3.2 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,

$$S_\rho = \frac{1}{2}\int (f_1^{1/2} - f_2^{1/2})^2 dx = \frac{1}{2}\int (1 - \frac{f_2^{1/2}}{f_1^{1/2}})^2 dF_1(x),$$

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 cryptocurrency 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). Block 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 cryptocurrency 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 cryptocurrency 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 cryptocurrencies 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 cryptocurrencies and stocks may mainly caused by the changes of stocks' distributions. We will dive deeper on this part in Section 4.

Table 3.2: Test equality of univariate densities: cryptocurrencies & 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 3.3: 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, **** $0.01 \leq p < 0.05$, *and ** $0.05 \leq p < 0.10$ *under the null of independence of returns.*

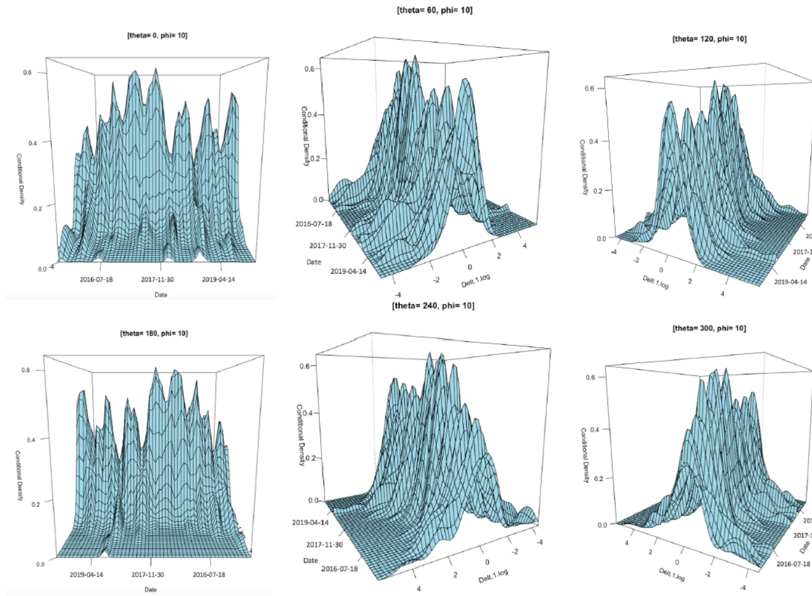


Figure 3.2: Density of NASDAQ: pre-COVID

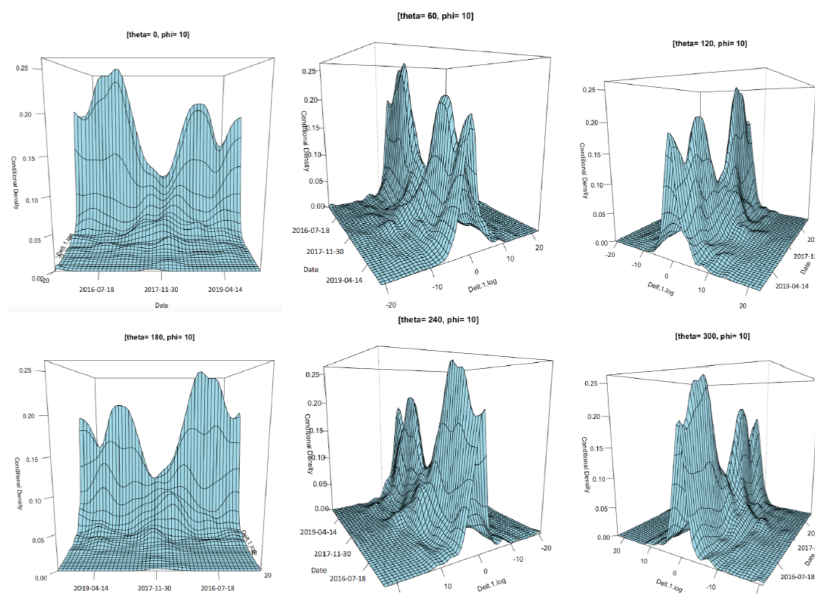


Figure 3.3: Density of Bitcoin: pre-COVID

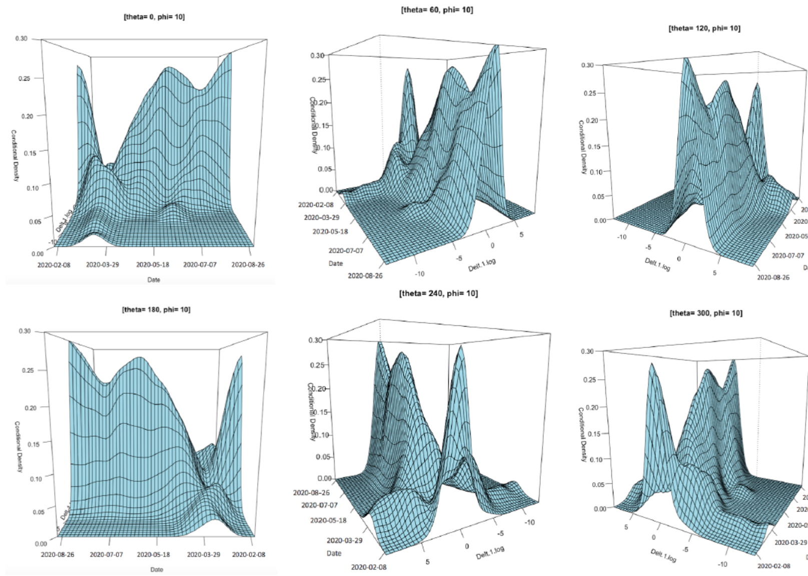


Figure 3.4: Density of NASDAQ: COVID era

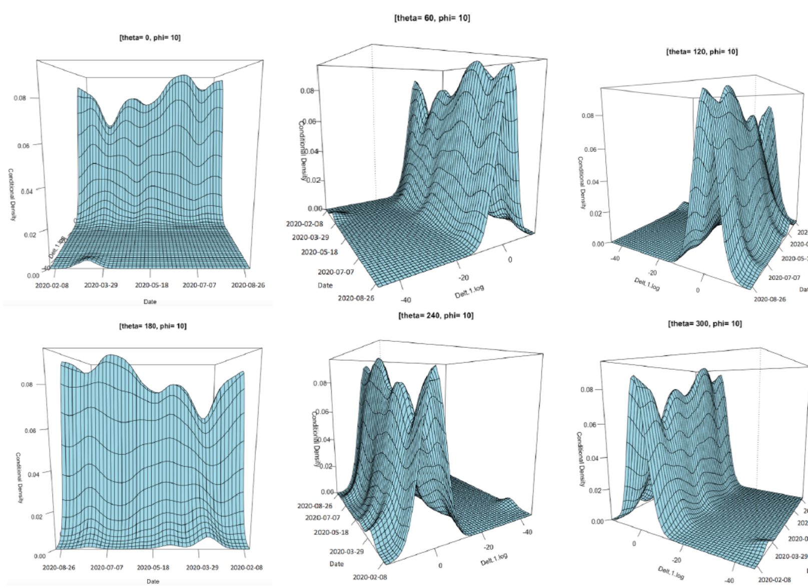


Figure 3.5: Density of Bitcoin: COVID era

3.3.3 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 [3]. 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 $t+1$. 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

[3] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data$_{L}ibrary/det_{3}0_{i}nd_{p}ort.html$

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 cryptocurrencies' 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 cryptocurrencies' (shown in Table 3).

Table 3.4: 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*, *** $0.01 \leq p < 0.05$*, and ** $0.05 \leq p < 0.10$ *under the null of independence of returns.*

3.3.4 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 cryptocurrencies for portfolio diversification.

Table 3.5: 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, **** $0.01 \leq p < 0.05$, *and ** $0.05 \leq p < 0.10$ *under the null of independence of returns.*

3.4 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 cryptocurrency 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.

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

where $S_\rho(f_{s_i,t_2}, f_{c_j,t_2})$ stands for the entropy metric between stock i and cryptocurrency 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 cryptocurrency 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 cryptocurrency 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 cryptocurrencies 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:

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

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 3.6: DID decomposition

Distribution	Stock	cryptocurrency	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	$S_\rho(f_{s_i,t_2}, f_{s_i,t_1})$	$S_\rho(f_{c_j,t_2}, f_{c_j,t_1})$	$S_\rho(f_{s_i,t_2}, f_{c_j,t_2}) - S_\rho(f_{s_i,t_1}, f_{c_j,t_1})$

3.5 Three-Period Analysis and the Vaccine Effect

Previously, we have confirmed that NASDAQ daily return has the most similar density and co-dependence with Bitcoin daily return, generally. However, the after COVID-19 outbreak in early 2020, even S&P500 daily return density is statistically closely dependent on, and indifferent from Bitcoin daily return. However, it includes only data up to 2020, when the COVID-19 pandemic occurred, and excludes 2021, when a vaccine is being developed and the global vaccination rate is increasing. Therefore, it is important to analysis the density similarity between cryptocurrencies and other assets after the vaccine rollout. In this section, we would like to discover the effect of COVID-19 outbreak as well as the vaccine rollout on density similarity between stocks and cryptocurrencies.

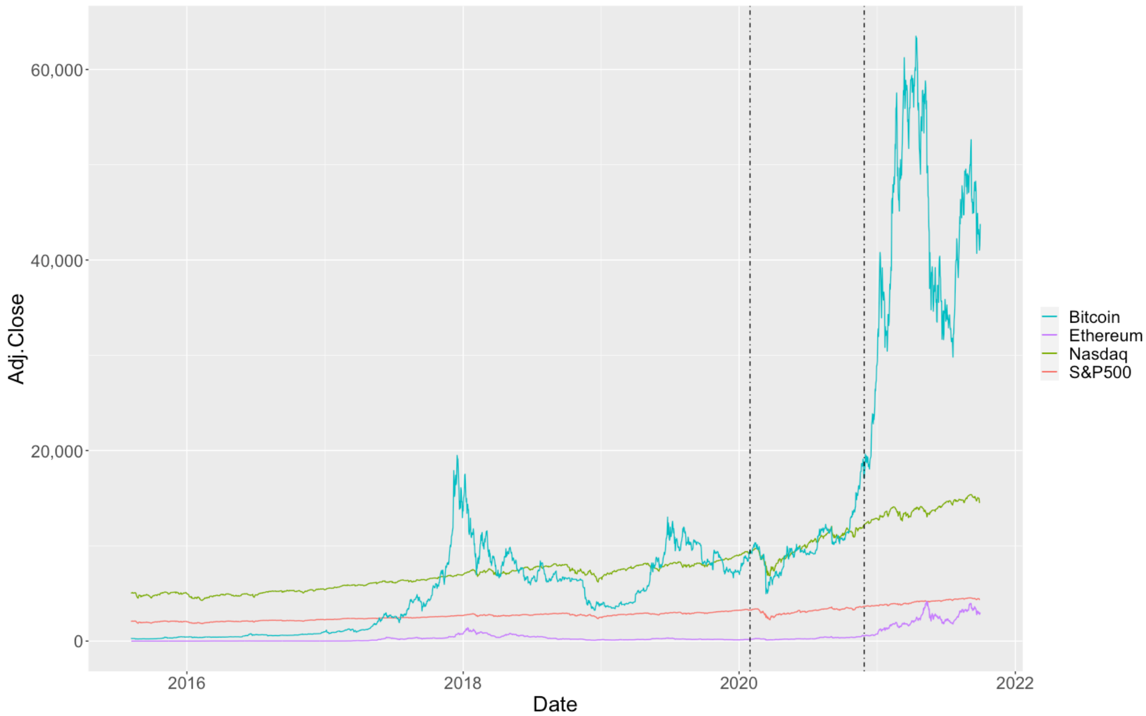
The cryptocurrency data and stock market index data set consists of daily spot exchange rates in units of US dollars are from Yahoo Finance. The price observations of Bitcoin (BTC-USD), Ethereum (ETH-USD), S&P500 stock market index (ˆGSPC) and NASDAQ stock market index (ˆIXIC) range from 6 August 2015 to 30 September 2021. We conduct our analysis across three time periods: Pre-COVID (August 6th , 2015 – January 31st, 2020), COVID era (February 1st, 2020 – November 30th, 2020) and Vaccine era (December 1st, 2020 – September 30th, 2021).

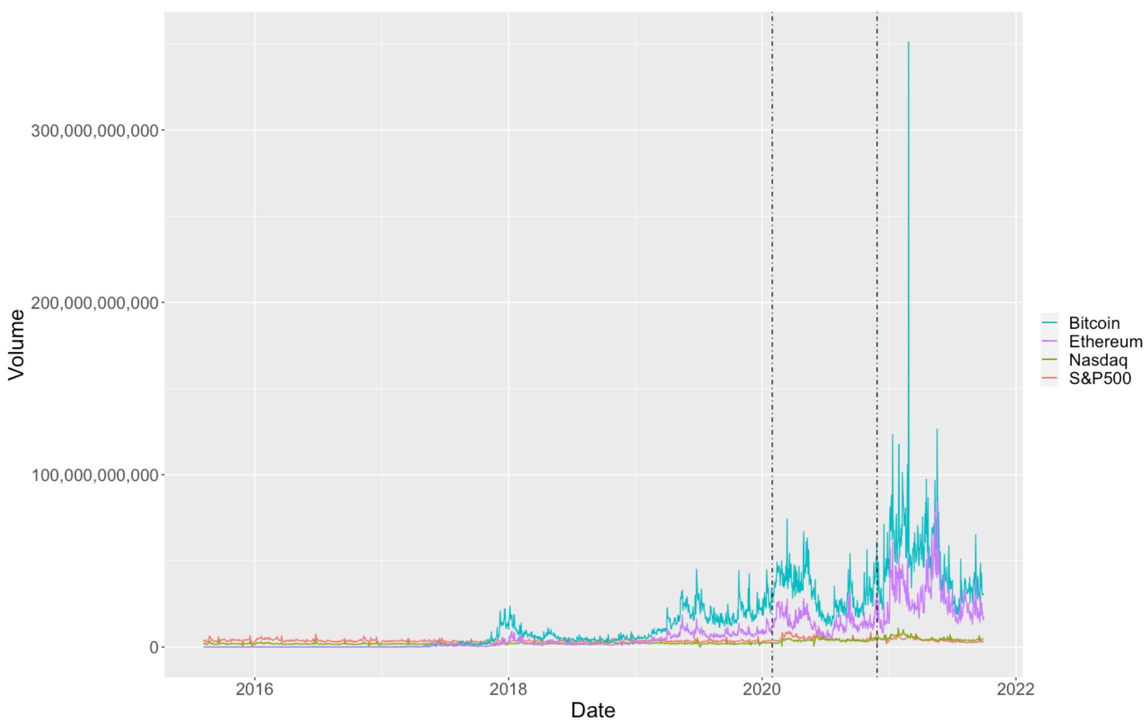
Firstly, we document main statistical properties of time series for the returns of S&P500 stock market index, NASDAQ stock market index, Bitcoin and Ethereum Figure 1 illustrates the time evolution of prices, volumes and daily log-returns for S&P500, NASDAQ, Bitcoin and Ethereum. We observe that both stock and cryptocurrency prices increased along the time, and the increases of cryptocurrency prices were relatively profound after COVID outbreak rather than the increase of stock prices. Notably, the Bitcoin price soared dramatically after the vaccine rollout. In addition, we notice that the volume of cryptocurrencies also increased after COVID outbreak, and the volume of cryptocurrencies became much more volatile simultaneously. Another thing that is worth to mention is the daily log-returns of both stock and cryptocurrency became more volatile after COVID outbreak, while after the roll-out of vaccines, the stock returns turned to be stable while the cryptocurrency returns were still extremely volatile. This highlighted the difficulty to predict cryptocurrency performance especially after COVID outbreak, thus demystifying the density similarity between stock and cryptocurrency returns is important for both scholars and investors.

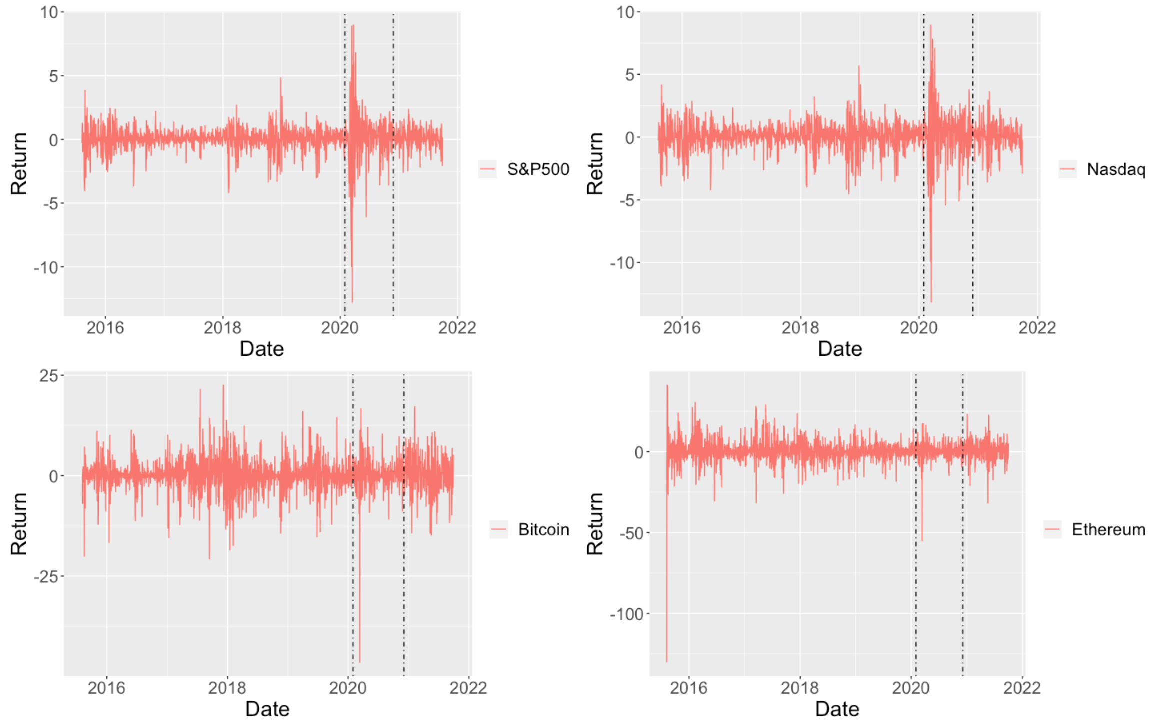
The descriptive statistics of daily log-returns are reported in Table 1. Along all the three periods, the daily returns of cryptocurrency markets exhibit high variability and excess kurtosis comparing with stock markets, especially the cryptocurrency returns had extremely high kurtosis during the COVID era. We observed that the stock market was more volatile than ever during the COVID era, and after the effective vaccine rollout, the stock market became less volatile and had a similar volatility as pre-COVID period. Overall, the daily returns of SP500, Bitcoin and Ethereum increased along the time, while the daily return of NASDAQ arrived that its peak during COVID era, and dropped afterwards.

The deviations from the Normal distribution are confirmed by the Jarque–Bera test that rejects the null hypothesis of normality. 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 cryptocurrency market is not as efficient as stock or foreign exchange markets, which display a complete lack of predictability (Lahmiri and Bekiros, 2018).

[insert table 3.7 here]







In this section, we still employ Maasoumi and Racine (2002) to test the equality of univariate densities between stock and cryptocurrency daily log-returns. In our method, we 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 (Granger et al., 2004) for testing the null of equality of two univariate density (or probability) functions. For continuous variables,

$$S_\rho = \frac{1}{2}\int (f_1^{1/2} - f_2^{1/2})^2 dx = \frac{1}{2}\int (1 - \frac{f_2^{1/2}}{f_1^{1/2}})^2 dF_1(x),$$

where $f_1 = f(x)$ and $f_2 = f(y)$ are the marginal densities of the random variables X and Y.

We estimate the metric S_ρ for the daily returns data for $x = Retur_{crypto}$ 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 and 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 $y = NASDAQ$ among all the three 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. During COVID era, S&P500 returns distribution is also statistically closely dependent on, and indifferent from Bitcoin's.

Comparing S_ρ among all the three time periods, we conclude that the values of S_ρ decrease generally in all cases, sometimes dramatically during COVID era, while the values of S_ρ increase (and even higher than pre-COVID period) in all cases during vaccine era. This suggests that the densities of cryptocurrency and stock index returns became more similar with the advent of COVID-19, while the densities become less similar after the vaccine rollout. This is consistent with our intuitive assumption: highly volatile assets like cryptocurrencies behave more similarly to other assets during down turns, compared to upturns.

In summary, the cryptocurrency market has recently increased in complexity and thus unpredictability is increasing due to the development of a vaccine, the complexity of the cryptocurrency market has still been increasing since then. Vaccination rates are rising worldwide, but it is estimated that many other factors, such as the emergence of the delta virus or foreshadowing of tapering, are adding to the complexity of the cryptocurrency market.

[insert table 3.8 here]

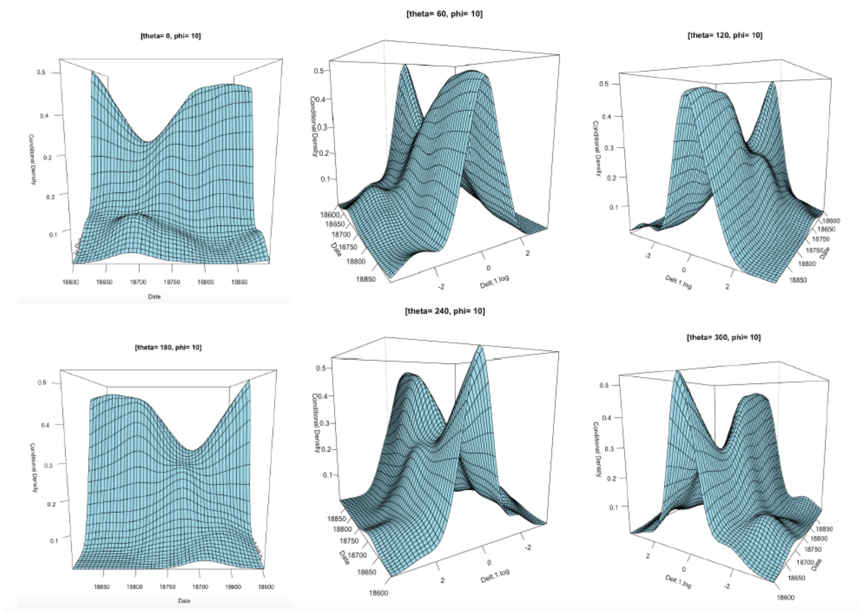


Figure 3.6: Density of NASDAQ: Vaccine era

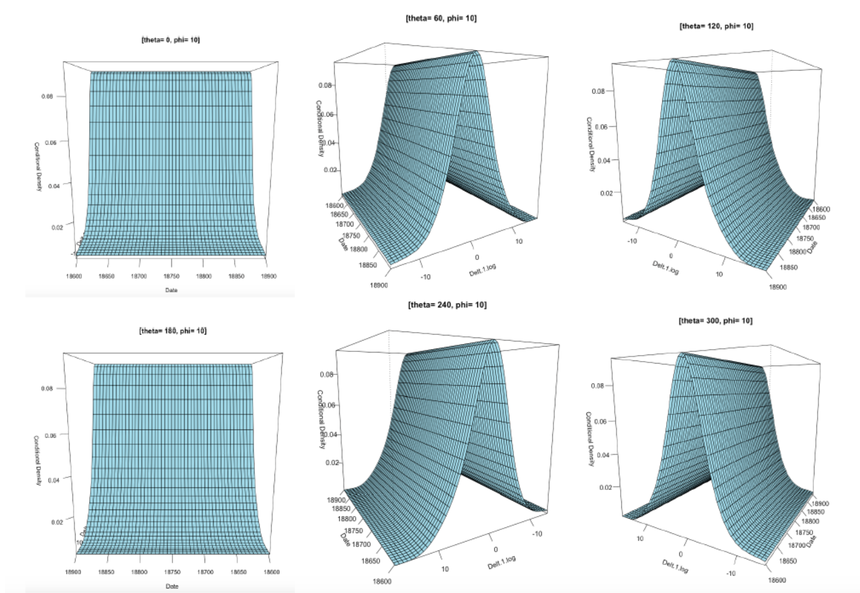


Figure 3.7: Density of Bitcoin: Vaccine era

Table 3.7: Descriptive statistics for three-period analysis

	Pre-COVID (2015/08/06 – 2020/01/31)				COVID era (2020/02/01 – 2020/11/30)				Vaccine era (2020/12/01 – 2021/09/30)			
Daily Log-Return	S&P500	NASDAQ	Bitcoin	Ethereum	S&P500	NASDAQ	Bitcoin	Ethereum	S&P500	NASDAQ	Bitcoin	Ethereum
Observations	1128	1128	1639	1638	208	208	296	296	208	208	303	303
Mean	0.04	0.05	0.21	0.25	0.05	0.13	0.22	0.37	0.08	0.08	0.28	0.54
Standard deviation	0.86	1.04	3.89	7.09	2.39	2.46	4.21	5.54	0.78	1.09	4.36	5.96
Skewness	-0.57	-0.51	-0.18	-3.44	-0.79	-0.91	-4.29	-3.29	-0.42	-0.43	-0.15	-0.51
Kurtosis	4.12	3.15	4.72	72.46	6.65	5.78	50.34	33.28	1.09	1.36	1.4	3.67
Augmented Dickey-Fuller (ADF)	-10.978 **	-11.258 **	-10.927 **	-10.926 **	-6.7324 **	-6.5787 **	-6.7698 **	-6.8565 **	-6.9732 **	-6.1272 **	-6.0104 **	-5.9225 **
Jarque–Bera	862.5 ***	518.27 ***	1538.9 ***	362524 ***	416.4 ***	327.18 ***	33078 ***	14674 ***	17.136 ***	23.487 ***	27.155 ***	187 ***

*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.*

Table 3.8: Test equality of univariate densities: cryptocurrencies and stocks for three-period analysis

	Pre-COVID (2015/08/06 – 2020/01/31)		COVID era (2020/02/01 – 2020/11/30)			Vaccine era (2020/12/01 – 2021/09/30)		
Daily Log-Return	S_ρ	p-value	S_ρ	p-value	delta	S_ρ	p-value	delta
S&P500 and Bitcoin	0.1988	<2.22e-16 ***	0.0350	0.010101 *	-0.16382064	0.3695	<2.22e-16 ***	0.1706796
S&P500 and Ethereum	0.3318	<2.22e-16 ***	0.0894	<2.22e-16 ***	-0.24237771	0.4612	<2.22e-16 ***	0.129419
NASDAQ and Bitcoin	0.1561	<2.22e-16 ***	0.0291	0.020202 *	-0.12703315	0.2738	<2.22e-16 ***	0.1176708
NASDAQ and Ethereum	0.2822	<2.22e-16 ***	0.0768	<2.22e-16 ***	-0.20537369	0.3699	<2.22e-16 ***	0.0876879

*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.*

3.6 Conclusion

This chapter 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 cryptocurrencies 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 chapter, 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 cryptocurrencies became more similar and less independent. These results are meaningful because we revealed the similarity and dependence structure between cryptocurrency and stock distributions. This can be useful in applying existing theories on stocks to cryptocurrencies.

As a supplement of our previous study, we also examine newer data as we have observe the effective vaccines rollout, stock market volatility and the cryptocurrency prices peak to new high in 2021. By conducting a three-period analysis, we were trying to answer the question that whether the densities of cryptocurrencies and stock marker indexes became more / less similar after vaccination. we discovered that the cryptocurrency market has recently increased in complexity and thus unpredictability is increasing due to the development of a vaccine. We believe the examination of newer data will drive more promising and effective policy implications.

Bibliography

[1] Al-Khazali, Osamah, Bouri Elie, and David Roubaud. "The impact of positive and negative macroeconomic news surprises: Gold versus Bitcoin." Economics Bulletin 38, no. 1 (2018): 373-382.

[2] Balcilar, Mehmet, Elie Bouri, Rangan Gupta, and David Roubaud. "Can volume predict Bitcoin returns and volatility? A quantiles-based approach." Economic Modelling 64 (2017): 74-81.

[3] Balcombe, Kelvin, and Iain Fraser. "Do bubbles have an explosive signature in markov switching models?." Economic Modelling 66 (2017): 81-100.

[4] Bariviera, Aurelio F. "The inefficiency of Bitcoin revisited: A dynamic approach." Economics Letters 161 (2017): 1-4.

[5] Bariviera, Aurelio F., María José Basgall, Waldo Hasperué, and Marcelo Naiouf. "Some stylized facts of the Bitcoin market." Physica A: Statistical Mechanics and its Applications484 (2017): 82-90.

[6] Baur, Dirk G., Thomas Dimpfl, and Konstantin Kuck. "Bitcoin, gold and the US dollar–A replication and extension." Finance research letters 25 (2018): 103-110.

[7] Begušić, Stjepan, Zvonko Kostanjčar, H. Eugene Stanley, and Boris Podobnik. "Scaling properties of extreme price fluctuations in Bitcoin markets." Physica A: Statistical Mechanics and its Applications 510 (2018): 400-406.

[8] Bouri, Elie, Georges Azzi, and Anne Haubo Dyhrberg. "On the return-volatility relationship in the Bitcoin market around the price crash of 2013." Economics 11, no. 1 (2017).

[9] Bouri, Elie, Naji Jalkh, Peter Molnár, and David Roubaud. "Bitcoin for energy commodities before and after the December 2013 crash: diversifier, hedge or safe haven?." Applied Economics 49, no. 50 (2017): 5063-5073.

[10] Bouri, Elie, Peter Molnár, Georges Azzi, David Roubaud, and Lars Ivar Hagfors. "On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier?." Finance Research Letters 20 (2017): 192-198.

[11] Bouri, Elie, Peter Molnár, Georges Azzi, David Roubaud, and Lars Ivar Hagfors. "On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier?." Finance Research Letters 20 (2017): 192-198.

[12] Bouri, Elie, Rangan Gupta, Aviral Kumar Tiwari, and David Roubaud. "Does Bitcoin hedge global uncertainty? Evidence from wavelet-based quantile-in-quantile regressions." Finance Research Letters 23 (2017): 87-95.

[13] Bouri, Elie, Syed Jawad Hussain Shahzad, and David Roubaud. "Co-explosivity in the cryptocurrency market." Finance Research Letters 29 (2019): 178-183.

[14] Card, David, and Alan B. Krueger. "Minimum wages and employment: A case study of the fast food industry in New Jersey and Pennsylvania." (1993).

[15] Chu, Jeffrey, Stephen Chan, Saralees Nadarajah, and Joerg Osterrieder. "GARCH modelling of cryptocurrencies." Journal of Risk and Financial Management 10, no. 4 (2017): 17.

[16] David, S. A., C. M. C. Inácio Jr, and José A. Tenreiro Machado. "The recovery

of global stock markets indices after impacts due to pandemics." Research in International Business and Finance 55 (2021): 101335.

[17] Del Giudice, Vincenzo, Pierfrancesco De Paola, and Francesco Paolo Del Giudice. "COVID-19 infects real estate markets: Short and mid-run effects on housing prices in Campania region (Italy)." Social sciences 9, no. 7 (2020): 114.

[18] Dyhrberg, Anne Haubo. "Bitcoin, gold and the dollar–A GARCH volatility analysis." Finance Research Letters 16 (2016): 85-92.

[19] Gherghina, Ștefan Cristian, Daniel Ștefan Armeanu, and Camelia Cătălina Joldeș. "Stock market reactions to Covid-19 pandemic outbreak: Quantitative evidence from ARDL bounds tests and Granger causality analysis." International journal of environmental research and public health 17, no. 18 (2020): 6729.

[20] Giannerini, Simone, Esfandiar Maasoumi, and Estela Bee Dagum. "Entropy testing for nonlinear serial dependence in time series." Biometrika 102, no. 3 (2015): 661-675.

[21] Granger, Clive WJ, and Namwon Hyung. "Occasional structural breaks and long memory with an application to the SP 500 absolute stock returns." Journal of empirical finance11, no. 3 (2004): 399-421.

[22] Granger, Clive WJ, Bwo-Nung Huangb, and Chin-Wei Yang. "A bivariate causality between stock prices and exchange rates: evidence from recent Asianflu." The Quarterly Review of Economics and Finance 40, no. 3 (2000): 337-354.

[23] Gulko, Les. "The entropy theory of stock option pricing." International Journal of Theoretical and Applied Finance 2, no. 03 (1999): 331-355.

[24] Katsiampa, Paraskevi. "Volatility estimation for Bitcoin: A comparison of GARCH models." Economics Letters 158 (2017): 3-6.

[25] Kim, Kyungwon, and Minhyuk Lee. "The Impact of the COVID-19 Pandemic on the Unpredictable Dynamics of the Cryptocurrency Market." Entropy 23, no. 9 (2021): 1234.

[26] Kristoufek, Ladislav. "What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis." PloS one10, no. 4 (2015): e0123923.

[27] Kullback, Solomon, and Richard A. Leibler. "On information and sufficiency." The annals of mathematical statistics 22, no. 1 (1951): 79-86.

[28] Lahmiri, Salim, and Stelios Bekiros. "Chaos, randomness and multi-fractality in Bitcoin market." Chaos, solitons fractals106 (2018): 28-34.

[29] Lahmiri, Salim, and Stelios Bekiros. "Cryptocurrency forecasting with deep learning chaotic neural networks." Chaos, Solitons Fractals 118 (2019): 35-40.

[30] Lahmiri, Salim, Stelios Bekiros, and Antonio Salvi. "Long-range memory, distributional variation and randomness of bitcoin volatility." Chaos, Solitons Fractals 107 (2018): 43-48.

[31] Li, Xin, and Chong Alex Wang. "The technology and economic determinants of cryptocurrency exchange rates: The case of Bitcoin." Decision support systems 95 (2017): 49-60.

[32] Maasoumi, Esfandiar, and Jeff Racine. "Entropy and predictability of stock market returns." Journal of Econometrics107, no. 1-2 (2002): 291-312.

[33] Naeem, Muhammad Abubakr, Elie Bouri, Zhe Peng, Syed Jawad Hussain Shahzad, and Xuan Vinh Vo. "Asymmetric efficiency of cryptocurrencies during COVID19." Physica A: Statistical Mechanics and its Applications 565 (2021): 125562.

[34] Naeem, Muhammad Abubakr, Saba Qureshi, Mobeen Ur Rehman, and Faruk Balli. "COVID-19 and cryptocurrency market: Evidence from quantile connectedness." Applied Economics (2021): 1-27.

[35] Nakayama, Ken, and Gerald H. Silverman. "Serial and parallel processing of visual feature conjunctions." Nature 320, no. 6059 (1986): 264-265.

[36] Osterrieder, Joerg, and Julian Lorenz. "A statistical risk assessment of Bitcoin and its extreme tail behavior." Annals of Financial Economics 12, no. 01 (2017): 1750003.

[37] Pichl, Lukáš, and Taisei Kaizoji. "Volatility analysis of bitcoin." Quantitative Finance and Economics 1, no. 4 (2017): 474-485.

[38] Rényi, Alfréd. "On measures of entropy and information." In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics, vol. 4, pp. 547-562. University of California Press, 1961.

[39] Renyi, Alfred. "On the number of endpoints of a k-tree." Studia Sci. Math. Hungar 5 (1970): 5-10.

[40] Sarkodie, Samuel Asumadu, Maruf Yakubu Ahmed, and Phebe Asantewaa Owusu. "COVID-19 pandemic improves market signals of cryptocurrencies–evidence from Bitcoin, Bitcoin Cash, Ethereum, and Litecoin." Finance Research Letters 44 (2022): 102049.

[41] Shannon, Claude Elwood. "A mathematical theory of communication." The Bell system technical journal 27, no. 3 (1948): 379-423.

[42] Skaug, Hans Julius, and Dag Tjøstheim. "Testing for serial independence using

measures of distance between densities." In Athens Conference on Applied Probability and Time Series Analysis, pp. 363-377. Springer, New York, NY, 1996.

[43] Stavroyiannis, Stavros. "Value-at-risk and related measures for the Bitcoin." The Journal of Risk Finance (2018).

[44] Thies, Sven, and Peter Molnár. "Bayesian change point analysis of Bitcoin returns." Finance Research Letters 27 (2018): 223-227.

[45] Vukovic, Darko, Moinak Maiti, Zoran Grubisic, Elena M. Grigorieva, and Michael Frömmel. "COVID-19 pandemic: Is the crypto market a safe haven? The impact of the first wave." Sustainability 13, no. 15 (2021): 8578.

Chapter 4

Do Cryptocurrencies and Other Assets Converge? A Clustering Analysis of Asset Returns

Since the first decentralized cryptocurrency, Bitcoin, was created in the year of 2009, cryptocurrencies gain a growing attention from the media, academics, and finance industry. However, this new market is very diverse and difficult to predict. In this chapter, we aim to examine the prospects for clustering, or convergence of asset classes. In the first instance, we examine if a set of cryptocurrencies form identifiable clusters within this class. Using entropy metric to assess "similarity" of entire distributions, we implement Agglomerative Hierarchical Clustering technique to examine whether or not cryptocurrencies are converging to "clubs" with similar distributions of returns. To arrive at a more convincing conclusion, we also apply the K-means Clustering to justify our results. We discover that cryptocurrencies share similar geographic locations and similar functions tend to converge to same clusters. We also observe another potential explanation to our results called the "Coinbase effect". In the second stage, we examine if these clusters include other asset classes, such as commodities. We find

cryptocurrencies and commodities are separated into different clusters using entropy metric as cluster proximity, which is consistent with intuitive assumptions. We also find that the cluster that contains the distributions of Coal (COAL) and Petroleum and Natural Gas (OIL) have smaller distance to cryptocurrency distributions. To conclude, our work will help to enhance the profiling of the clusters with additional insights. As a result, this work offers a description of the market and a methodology that can be reproduced by investors that want to understand the main trends on the market and that look for cryptocurrencies with different financial performance.

Keywords:

Financial markets, Cryptocurrency, Bitcoin, Convergence, Entropy, Clustering analysis, Agglomerative Hierarchical Clustering, K-means Clustering

4.1 Introduction

The cryptocurrency market consists of more than 4,000 cryptocoins with over 800 trades per second and more than 280 exchanges. Since the first peer-to-peer and decentralised digital currency, Bitcoin, was fisrt created in 2008 and mined in 2009, cryptocurrencies has become a huge new market in a very short term (Nakamoto, 2009). While cryptocurrencies were originally intended to enable anonymous wire transfers and online purchases, they have indeed become a powerful investment tool nowadays.

However, this new market is very diverse. Cryptocurrencies with different technologies, purposes and user base coexist and form a highly heterogeneous market that is difficult to understand and to manage for those addressing a good investment allocation.

As other traditional assets, the value of cryptocurrencies swing based on news events, but cryptocurrencies have no physical assets or governments to back their value. Moreover, the cryptocurrency market is new, based on a still developing technology, highly speculative and small in comparison to others. As a result, it is highly volatile with big upswings, bubbles, and sudden market downturns. Being a market so novel, big, diverse and volatile, it needs to be understood. Investment managers are constantly challenged to identify assets and asset classes that are, in some sense, orthogonal to existing assets, so as to decide if they should be included in further diversification of assets under management.

Several categorization efforts have been made so far. Burniske and Tatar (2017) classified over 200 cryptocurrencies into three classes of assets included capital asset, consumable/transformable assets and store of value asset. However, this classification is highly subjective and only cover a small fraction of the cryptocurrencies. Another approach leverages on statistical methods to analyze the financial performance of cryptocurrencies. Stosic et al. (2018) analyze cross correlations between price changes of different cryptocurrencies using methods of random matrix theory and minimum spanning trees, discover distinct community structures in their minimum spanning trees. Hu et al. (2019) present stylized facts on the asset pricing properties of cryptocurrencies by summarizing statistics on cryptocurrency return properties and measures of common variation for secondary market returns on 222 digital coins, and find a large degree of skewness and volatility in the population of returns. Song et al. (2019) analyze the structure of the cryptocurrency market based on the correlation-based agglomerative hierarchical clustering and minimum spanning tree. They propose a Bitcoin-Ethereum filtering mechanism, and discover the leadership of the Bitcoin and Ethereum in the market and six homogeneous clusters composed of relatively less-traded cryptocurrencies. Sigaki et al.(2019) use time series to represent the cryptocurrencies, use permutation entropy and statistical complexity over sliding time-windows of price log returns to quantify the dynamic efficiency of 437 time series of cryptocurrencies. Pele et al. (2020) separate cryptocurrencies from the classical assets by using various classification techniques applied to the daily time series of log-returns, mainly due to their tail behavior. All these approaches reveal that it is possible to establish different groups of cryptocurrencies in terms of their financial performance. And identifying them, it is useful to better understand the cryptocurrency market, but also for building a diversified portfolio. Therefore, it is important for us to study the clustering behavior of cryptocurrency market.

Literature also focused on clustering of cryptocurrencies. Song et al. (2019) applied the classical methodology based on MST algorithms (Mantegna, 1999) to filter out the influence of Bitcoins and Ethereum, and it detects six homogeneous clusters. However, the structure found does not remain stable after the announcement of regulations from various countries. Zieba et al. (2019) use clustering method together with other methods, such as VAR models and Granger causality tests, to find that Bitcoin shock prices are not transmitted to the prices of other cryptocurrencies, thus suggested that Bitcoin shall not be generalized to all cryptocurrencies. Corbet et al. (2018) show that cryptocurrencies are highly connected among themselves and disconnected from mainstream assets, such as bonds, stocks, S&P500 and gold. All these articles evidence the complexity of the underlying structure in the cryptocurrency market, where some cryptocurrencies influence others even in unexpected ways.

More recently, an increasing body of the literature studies their role in portfolio diversification. For instance, Platanakis and Urquhart (2020) examine the out-of-sample benefits of including Bitcoin in a stock-bond portfolio. The empirical findings, using weekly data from October 2011 to June 2018, indicate improvement in different measures of risk-adjusted returns and are robust to a variety of asset allocation strategies, risk aversion levels, rolling windows, the inclusion of transaction costs and different assets, allowance for short selling and optimization techniques. Liu (2019) apply different portfolio selection models, conducting a large set of robustness checks to analyze portfolio performance considering 10 cryptocurrencies. They point out that diversification among cryptocurrencies can enhance performance under the Sharpe ratio and utility criteria. However, optimal portfolios are often outperformed by the naive equal weights portfolio and estimation error in parameters of returns distribution may offset gains of diversification. Kajtazi and Moro (2019) use a mean-CVar approach to study bitcoin diversification role in different markets, USA, China, and Europe, finding converging results across markets for performance improvement, more due to increase in returns rather than reduction of risk. These conclusions holdmainly for long-only and naïve portfolios, in conformity to other findings in the literature (DeMiguel et al., 2009; Platanakis et al., 2018). Under a different approach, Guesmi et al. (2019) use a set of time series models to analyze joint dynamics of Bitcoin and different financial assets and conclude that investors may benefit from hedging and diversification gains adding bitcoin to their portfolios.

Cryptocurrencies are examined with this question in mind, in a large set of comparisons with other assets. These comparisons may be viewed as an assessment of "similarity" with existing asset classes. In this paper, we hope to shed light on these issues in several important ways that respond to the limitations discussed above. First, following Maasoumi et al. (2007) and Maasoumi & Wang (2008), we propose a new concept of convergence based on the similarity of the distributions of asset daily log-returns. The similarity is measured by the normalization of the Bahattacharya-Matusita-Hellinger Entropy measure proposed in Granger et al. (2004). This entropy measure goes beyond the first and second moments of the distributions and is able to summarize the information of the "entire" distribution. Since it is also a "metric" measure of the "distance" between two distributions (as contrasted with "divergence"), our method provides a detailed picture of the changes in the distance of the distributions of asset daily log-returns. Our prior work (Maasoumi & Wu, 2021) is based on the most comprehensive criterion for assessment of such similarity, the similarity of the entire distributions. We employed metric entropy testing to rigorously infer equality of distributions and co-dependence.

In this chapter, we examine the prospects for clustering, or convergence of asset classes. In the first instance, we examine if a set of cryptocurrencies form identifiable clusters within this class. In the second stage, we examine if these clusters include other asset classes, such as commodities. Many analysts have advocated that these currencies be viewed as commodities, because of their price behavior in recent markets. The clustering method we employ is similarly comprehensive in its definition of cluster membership. This is based on the entire statistical distribution of an asset return as the probability law that generates it. In this sense, testing for equality of asset distributions is a technique that is shared with our prior work.

The rest of the chapter is organized as follows: Section 2 presents the methodology that we used in this chapter. We conducted clustering analysis depends on measures of proximity or similarity. Section 3 described that data we leveraged on.

Section 4 discussed our empirical results. In details, we conduct the cluster analysis of cryptocurrencies, cluster analysis of cryptocurrencies and commodities, as well as comparing our results using Agglomerative Hierarchical Clustering and K-means Clustering. Section 5 concluded our findings in this chapter.

4.2 Empirical Methodology

4.2.1 Entropy Measures of Distributional Distance

In our chapter, we define the convergence as the similarity of the distributions of daily log returns of two cryptocurrencies. Mathematically, two cryptocurrencies are converging if

$$H_0 : f_i = f_j,$$

where f_i and f_j are marginal densities of the daily log returns of cryptocurrency i and cryptocurrency j in our analysis. The traditional moment relations can test the equality of distributions by either mean or variance, but this kind of measurement is problematic because it lack the power to contain all information in distributions. While our novel metric entropy measures can compare the entire distributions of cryptocurrency returns, which capture all information contains in distributions. Another highlight of our metric entropy measures is that it is a *metric*, which means it satisfies the triangularity rule of distances. Hence, it measures not only divergence but also distance between two distributions, which will be especially suitable when implementing cluster analysis, and go beyond hypothesis tests of equality.

We utilize a metric entropy measure S_ρ following Maasoumi & Racine (2002). This entropy is a normalization of the Bhattacharya-Matusita-Hellinger measure of distance (Granger et al., 2004), and provides a quantified distance between distributions of cryptocurrency returns. It is given by

$$S_\rho = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(f_i^{\frac{1}{2}} - f_j^{\frac{1}{2}})^2 dxdy.$$

In other words, our null hypothesis can be written as:

$$H_0 : S_\rho = 0.$$

Thus, the distributions of two cryptocurrencies converge if and only if we cannot reject the null hypothesis in our analysis.

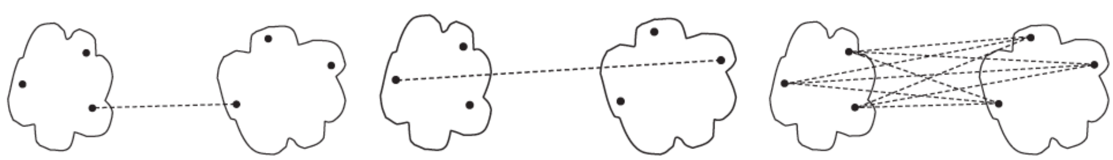
The following properties are satisfied by this entropy measure S_ρ: (i) S_ρ is well defined for both continuous and discrete variables; (ii) S_ρ is normalized to zero if X and Y are independent, and lies between 0 and 1; (iii) S_ρ is metric, which means it is a true measure of distance and not just of Kullback-Leibler divergence; (iv) The entropy measure S_ρ is invariant under continuous and strictly increasing transformations $h()$. We employ block bootstrap resampling techniques to conduct tests based on kernel density implementations of our statistic, with cross validation determining optimal bandwidths.

4.2.2 Cluster Analysis

Agglomerative Hierarchical Clustering

Based on the entropy measure S_ρ in previous sections, we implement Agglomerative Hierarchical Clustering techniques to examine whether or not cryptocurrencies are converging to "clubs" with similar distributions of returns (Maasoumi & Wang, 2008).

The Agglomerative Hierarchical Clustering is the most common type of hierarchical clustering used to group objects in clusters based on their similarity. The algorithm Starts with the points as individual clusters, and at each step, merge the closest pair of clusters. This requires defining a notion of cluster proximity. In this chapter, we will apply three different calculations of the proximity between two clusters: complete linkage, single linkage and average linkage. In the complete linkage, the distance between two clusters is defined as the maximum value of all pairwise distances between the elements in cluster 1 and the elements in cluster 2, which produces more compact clusters; In single linkage, the distance between two clusters is defined as the minimum value of all pairwise distances between the elements in cluster 1 and the elements in cluster 2, which produces "loose" clusters; In average linkage, the distance between two clusters is defined as the average distance between the elements in cluster 1 and the elements in cluster 2. In Figure 1, we show the graph-based definitions of cluster proximity. The Agglomerative Hierarchical Clustering is often displayed graphically using a tree-like diagram called a dendrogram, which displays both the cluster-subcluster relationships and the order in which the clusters were merged.



(a) MIN (single link). (b) MAX (complete link). (c) Group average.

Figure 4.1: Graph-based definitions of cluster proximity

This chapter follows Hirschberg, Maasoumi & Slottje (2001), which employed clustering analysis based on the same entropy measure S_ρ to identify distinct dimensions in the 15 indicators of well-being that are commonly used. The algorithm is as follows:

1. Each cryptocurrency return variable is one cluster itself;
2. Compute the proximity matrix based on the same entropy measure S_ρ;

3. Merge the closest two clusters, and form a new cluster;

4. Repeat step 2-3 to update the proximity matrix to reflect the proximity between the new cluster and the original clusters;

5. Stop until only one cluster remains.

In our analysis, we use the complete linkage clustering algorithm, which finds the closest two groups based on the "farthest" observations between the two groups, as judged by a chosen criterion of distance. The complete linkage clustering algorithm requires a metric measure of distance, since any measures that violate the triangle rule will lead to inconsistent decisions (Hirschberg & Maasoumi, 2001). Therefore, our measurement of distance with entropy metric S_ρ guarantees the feasibility of this algorithm. As we mentioned earlier, one of advantages of Hierarchical Agglomerative Clustering method is it allows us to find the most suitable number of clusters. However, choosing the optimal number of clusters is a vital aspect of our analysis.

K-means Clustering

K-means clustering (MacQueen, 1967) is the most commonly used unsupervised machine learning algorithm for partitioning a given data set into a set of k clusters. The basic idea behind K-means clustering consists of defining clusters so that the total within-cluster variation is minimized, it clustering minimizes within-cluster variances, that is, squared Euclidean distances in our case, which makes the result easy to understand and interpret. The standard K-means algorithm is the Hartigan-Wong algorithm (1979), which defines the total within-cluster variation as the sum of squared distances Euclidean distances between items and the corresponding centroid:

$$W(C_k) = \sum_{x_i \in C_k} (x_i - \mu_k)^2,$$

where x_i is a data point belonging to the cluster C_k and μ_k is the mean value of the points assigned to the cluster C_k. Each observation x_i is assigned to a given cluster such that the sum of squares distance of the observation to their assigned cluster centers (μ_k) is minimized. The total within-cluster variation is defined as follows:

$$tot.withiness = \sum_{k=1}^{K} W(C_k) = \sum_{k=1}^{K} \sum_{x_i \in C_k} (x_i - \mu_k)^2,$$

The total within-cluster sum of square *tot.withiness* measures the compactness of the clustering and we want it to be as small as possible. K-means algorithm can be summarized as follows:

1. Specify the number of clusters (k) to be created using elbow method;

2. Select randomly k objects from the data set as the initial cluster centroids;

3. Assigns each observation to their closest centroid, based on the Euclidean distance between the object and the centroid;

4. For each of the k clusters, update the cluster centroid by calculating the new mean values of all the data points in the cluster. The centroid of a $K-th$ cluster is a vector of length p containing the means of all variables for the observations in the $K-th$ cluster; p is the number of variables.

5. Iteratively minimize the total within sum of square *tot.withiness*. In other words, iterate steps 3 and 4 until the cluster assignments stop changing or the maximum number of iterations is reached.

4.3 Data

The cryptocurrency data we used in this chapter is from CoinMarketCap [1]. This influential website for cryptocurrencies lists and ranks all major representatives of this asset class by providing a volume-per-exchange weighted market price, the market

[1] https://coinmarketcap.com

capitalisation and the total trading volume itself. We consider cryptocurrencies with top 50 highest market capitalization [2] Table 1 documented the name, abbreviation, market capitalization and volume of the cryptocurrencies with top 50 highest market capitalization. Our analysis used daily data and contained time period from January 1, 2019 to January 1, 2021. Therefore, we exclude 18 cryptocurrencies which have relative short histories, and use the remaining 32 cryptocurrencies for our analysis.

In each data set of cryptocurrencies, we have open price, intraday high price, intraday low price, close price (adjusted for splits), adjusted close price (adjusted for both dividends and splits), market capitalization and volume.

---

[2]The term market capitalization (or market cap) refers to a metric that measures the relative size of a cryptocurrency. It is calculated by: Market Cap = Current Price x Circulating Supply.

Table 4.1: Overview of cryptocurrencies for analysis

Rank	Name	Abbreviation	Market Cap ($)	Volume ($)	Exclude from Analysis
1	Bitcoin	BTC	611,351,681,172	55,093,310,438	
2	Ethereum	ETH	89,044,253,262	18,946,964,605	
3	Tether	USDT	21,228,581,016	76,275,962,035	
4	XRP	XRP	10,148,107,889	4,676,476,080	
5	Litecoin	LTC	9,156,541,209	8,898,953,873	
6	Polkadot	DOT	8,604,653,503	2,992,452,119	x
7	Bitcoin Cash	BCH	6,725,409,669	4,206,360,704	
8	Cardano	ADA	5,702,145,578	1,273,790,776	
9	Binance Coin	BNB	5,561,913,834	492,662,644	
10	Chainlink	LINK	4,850,105,588	1,390,965,360	
11	USD Coin	USDC	4,082,743,317	1,318,460,238	
12	Wrapped Bitcoin	WBTC	3,795,216,732	115,203,080	x
13	Bitcoin SV	BSV	3,089,652,572	388,116,627	
14	Stellar	XLM	2,832,570,834	417,164,429	
15	Monero	XMR	2,561,771,263	2,125,410,453	
16	EOS	EOS	2,507,061,549	2,531,463,330	
17	THETA	THETA	2,399,551,082	351,535,327	
18	TRON	TRX	1,941,458,934	864,302,641	
19	NEM	XEM	1,855,462,653	111,176,371	
20	Tezos	XTZ	1,525,665,155	202,102,105	
21	Crypto.com Coin	CRO	1,467,051,154	144,056,108	
22	Celsius	CEL	1,438,614,864	17,863,864	
23	UNUS SED LEO	LEO	1,355,397,178	10,115,030	x
24	VeChain	VET	1,352,923,475	255,606,413	
25	Uniswap	UNI	1,276,488,504	1,218,691,767	x
26	Dogecoin	DOGE	1,259,759,889	2,061,148,146	
27	Dai	DAI	1,219,604,894	420,889,163	x
28	Cosmos	ATOM	1,158,321,953	557,792,320	x
29	Binance USD	BUSD	1,125,383,198	1,630,986,398	x
30	Aave	AAVE	1,039,303,341	208,235,560	x
31	Synthetix	SNX	2,033,775,828	189,745,939	
32	Neo	NEO	1,601,725,575	679,868,891	
33	Compound	COMP	1,491,931,104	589,707,386	x
34	Maker	MKR	1,459,080,487	169,031,549	
35	SushiSwap	SUSHI	1,245,011,570	887,223,297	x
36	Huobi Token	HT	1,197,410,267	310,286,798	
37	Elrond	EGLD	1,190,289,362	217,917,378	x
38	IOTA	MIOTA	1,149,359,392	36,917,888	
39	FTX Token	FTT	1,086,314,790	29,742,332	x
40	Solana	SOL	1,059,078,241	30,739,714	x
41	Filecoin	FIL	1,048,187,194	173,466,067	x
42	Dash	DASH	1,037,213,333	421,676,095	
43	Revain	REV	991,918,528	7,497,311	
44	Zcash	ZEC	921,536,671	603,261,148	
45	yearn.finance	YFI	925,395,007	389,382,739	x
46	Avalanche	AVAX	908,573,876	109,696,557	x
47	Kusama	KSM	881,925,395	94,451,652	x
48	Ethereum Classic	ETC	878,465,983	951,148,462	
49	Algorand	ALGO	862,724,869	477,886,074	x
50	Decred	DCR	797,471,586	27,126,120	

To better compare performance between different cryptocurrencies, we calculate the logarithmic return of cryptocurrency i using its adjusted close price at time t:

$$r_{i,t} = 100 * [\ln(P_{i,t}) - \ln(P_{i,t-1})], \tag{4.1}$$

where $P_{i,t}$ denotes the adjusted close price of cryptocurrency i in USD at time t.

We now document main statistical properties of time series for the returns of cryptocurrencies in Table 3. The results in Table 3 shows the almost zero mean and small standard deviation for each crypto return series. Firstly, given that the risk-free rate r_f is set to be zero, the Sharpe ratio $S_p = (r_p - r_f)/\sigma_p$, a risk-adjusted return, is the mean return divided by its standard deviation. In the case of two cryptocurrencies with the same mean returns, a cryptocurrency with a higher Sharpe ratio can be considered as a superior investment asset. We notice that the cryptocurrencies with highest Sharpe ratios are: Celsius (CEL), Synthetix (SNX), THETA (THETA), Bitcoin (BTC), Chainlink (LINK), Huobi Token (HT) and Ethereum (ETH), which include the two cryptocurrencies with highest market cap (Bitcoin and Ethereum). Secondly, distribution is approximately symmetric if the skewness is within ± 0.5. In this sense, the return distributions of most cryptocurrencies are relatively asymmetric and left-skewed. Also, the kurtosis suggests that the distributions of all cryptocurrencies exhibit fat tails in comparison to that of the Gaussian distributions whose kurtosis is equal to three. To conclude, the returns of most cryptocurrencies exhibit high variability and excess kurtosis. Lastlty, we applied the Augmented-Dicker-Fuller (ADF) unit-root test of Dickey and Fuller (1979), 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. The test results show that the null hypothesis is rejected for all 32 crypto returns, which indicate the time series are stationary based on the strong statistical rejections on unit roots. The deviations from the Normal distribution are confirmed by the Jarque-Bera test that rejects the null hypothesis of normality. The test results show that all crypto returns in our analysis do not exhibit Normal distribution.

Figure 2 illustrates the time evolution of prices and returns for the 32 cryptocurrencies. From the plot of historical price data, we can observe that most crypto prices increased before the end of 2017, and arrived at their first peak of price at the end of 2017. Then the crypto prices decreased till the end of 2018, and soared dramatically to a new price peak since then. The cryptocurrencies exhibit this pattern include: ADA, BNB, BTC, DCR, DOGE, ETH, HT, LINK, LTC, NEO, SNX, THETA, TRX, VET, XEM, XLM and XMR. Some cryptocurrency prices fluctuated around their mean throughout the time period, including: BCH, DASH, ETC, MIOTA, MKR, USDC, USDT and ZEC. We also notice cryptocurrency CRO's price was stable before 2020, while increased before mid-2020 and then decreased. Also we found the prices of BSV and EOS decreased after 2019, the price of CEL increased all the time since it landed, the prices of XRP and XTZ dropped drastically at the end of 2020, and the price of REV kept decreasing since 2019. We also found that the series of cryptocurrency returns are stationary during the time period of our analysis, which also coincident with our ADF test results. And we also noticed that cryptocurrency returns peaked in March 2020. This was caused by the massive market crash on March 12-13 saw daily volumes hit $75.9 billion in a single day, the single greatest daily volume recorded in cryptocurrency history.

Table 4.2: Overview of commodities for analysis

Commodity	Abbreviation
Food Products	FOOD
Beer and Liquor	BEER
Tobacco Products	SMOKE
Recreation	GAMES
Printing and Publishing	BOOKS
Consumer Goods	HSHLD
Apparel	CLTHS
Healthcare, Medical Equipment, Pharmaceutical Products	HLTH
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
Restaurants, Hotels, Motels	MEALS
Banking, Insurance, Real Estate, Trading	FIN
Everything Else	OTHER

Notes: Data from Kenneth R. French 30 Industry Portfolios. The database assigns 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. (The database uses Compustat SIC codes for the fiscal year ending in calendar year $t-1$. Whenever Compustat SIC codes are not available, the database uses CRSP SIC codes for June of year t.) Then compute returns from July of t to June of $t+1$.

Table 4.3: Descriptive statistics for cryptocurrencies

Cryptocurrency	Obs	Mean	Std	Sharpe Ratio	Skewness	Kurtosis	ADF	Jarque-Bera
BTC	732	0.29	3.79	0.0765	-2.29	33.17	-8.4895 **	34396 ***
ETH	732	0.23	4.72	0.0487	-2.28	26.70	-8.4519 **	22517 ***
USDT	732	0.00	0.58	0.00	0.08	15.81	-12.456 **	7671.8 ***
XRP	732	-0.07	5.08	-0.0138	-1.35	29.45	-7.5929 **	26838 ***
LTC	732	0.20	5.02	0.0398	-0.67	11.16	-8.7664 **	3883.4 ***
BCH	732	0.10	5.46	0.0183	-1.08	20.10	-8.5571 **	12546 ***
ADA	732	0.20	5.31	0.0377	-1.02	11.87	-8.6426 **	4452.4 ***
BNB	732	0.25	4.70	-0.0532	-2.12	25.51	-9.264 **	20523 ***
LINK	732	0.51	6.67	0.0765	-0.20	14.58	-8.5009 **	6532.8 ***
USDC	732	0.00	0.55	0.00	0.46	12.60	-11.589 **	4900.6 ***
BSV	732	0.08	7.16	0.0112	2.91	44.19	-8.5081 **	60949 ***
XLM	732	0.01	5.14	0.0019	0.50	13.55	-8.5115 **	5664.4 ***
XMR	732	0.15	4.57	0.0328	-1.87	19.31	-9.0244 **	11877 ***
EOS	732	0.00	5.17	0.00	-1.42	14.89	-9.1991 **	7052.3 ***
THETA	732	0.53	6.78	0.0782	-0.65	10.43	-7.5369 **	3392.9 ***
TRX	732	0.05	5.19	0.0096	-1.45	15.14	-9.52 **	7298.2 ***
XEM	732	0.15	5.41	0.0277	0.44	5.45	-7.9895 **	937.39 ***
XTZ	732	0.20	5.96	0.0336	-1.04	15.68	-8.3612 **	7682.8 ***
CRO	732	0.14	6.54	0.0214	3.42	51.76	-8.3221 **	83605 ***
CEL	732	0.70	7.30	0.0959	-0.24	5.46	-8.9476 **	924.44 ***
VET	732	0.21	6.12	0.0343	-1.17	15.16	-8.392 **	7219.7 ***
DOGE	732	0.12	4.41	0.0272	1.03	20.84	-9.7068 **	13457 ***
SNX	732	0.74	8.21	0.0901	0.10	4.48	-8.4729 **	618.48 ***
NEO	732	0.09	5.24	0.0172	-0.80	10.32	-8.4938 **	3351 ***
MKR	732	0.03	5.72	0.0052	-3.51	58.30	-9.0152 **	105787 ***
HT	732	0.22	4.15	0.0530	-1.90	30.80	-8.5629 **	29547 ***
MIOTA	732	-0.03	5.09	-0.0059	-1.45	20.12	-8.9533 **	12682 ***
DASH	732	0.02	5.09	0.0039	-0.09	16.81	-8.3025 **	8677 ***
REV	732	-0.40	5.40	-0.0741	-3.37	63.58	-9.3197 **	125403 ***
ZEC	732	0.00	5.17	0.00	-0.88	8.34	-8.4652 **	2233.4 ***
ETC	732	0.02	4.95	0.0040	-1.53	17.98	-8.5885 **	10212 ***
DCR	732	0.12	4.93	0.0243	-1.47	16.76	-9.1469 **	8885.4 ***

*Notes: Std, S/R, ADF and JB are the abbreviations of the standard deviation, Sharpe ratio, Augmented Dickey Fuller and Jarque-Bera tests, respectively. 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 4.2: Time series plot for cryptocurrency prices and cryptocurrency returns (32 cryptocurrencies)

We also document descriptive statistics for commodities in Table 4 and plot the time series of commodity returns in Figure 3. Compared cryotos with commodities, a relative traditional asset class, we found that cryptocurrency returns exhibit significant high variability, excess skewness and kurtosis, as well as deviation from Normal distribution. These observations suggest that the cryptocurrency market is not as efficient as stock or foreign exchange markets, which display a complete lack of predictability (Lahmiri et al., 2018).

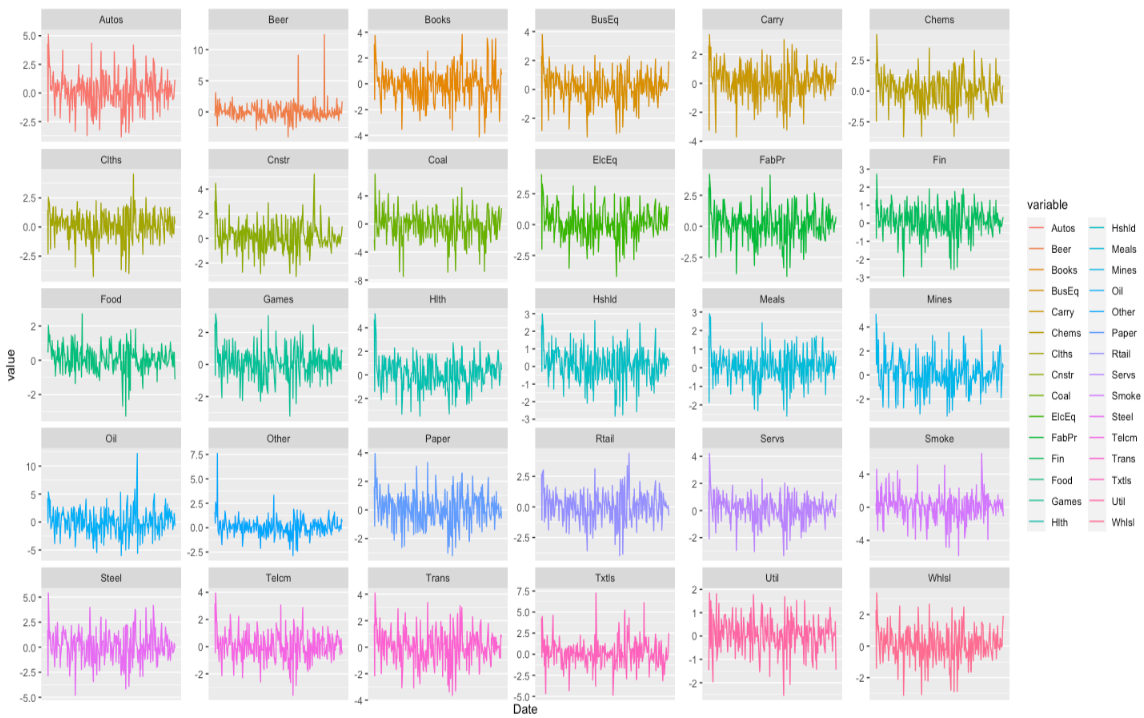


Figure 4.3: Time series plot for commodity returns (30 commodities)

Table 4.4: Descriptive statistics for commodities

Commodity	Obs	Mean	Std	Sharpe Ratio	Skewness	Kurtosis	ADF	Jarque-Bera
FOOD	253	0.07	0.74	0.0946	-0.34	2.16	-6.094 **	55.862 ***
BEER	253	0.05	1.43	0.0350	3.41	26.28	-5.5157 **	7908.6 ***
SMOKE	253	0.08	1.65	0.0485	-0.04	2.06	-5.2288 **	46.436 ***
GAMES	253	0.07	0.95	0.0737	0.08	0.97	-5.615 **	10.938 ***
BOOKS	253	-0.05	1.30	-0.0385	-0.06	0.89	-6.5994 **	9.1584 **
HSHLD	253	0.10	0.93	0.1075	-0.17	0.74	-6.2566 **	7.4648 **
CLTHS	253	0.10	1.24	0.0806	-0.45	1.11	-5.7245 **	22.205 ***
HLTH	253	0.12	1.26	0.0952	0.13	1.30	-5.6234 **	19.411 ***
CHEMS	253	0.08	1.25	0.0640	-0.03	0.69	-5.1162 **	5.4918 *
TXTLS	253	0.13	1.60	0.0813	0.61	2.95	-5.4978 **	110.48 ***
CNSTR	253	0.15	1.14	0.1316	0.47	1.89	-6.0044 **	48.46 ***
STEEL	253	0.07	1.52	0.0461	0.07	0.71	-5.8882 **	5.9409 *
FABPR	253	0.12	1.29	0.0930	-0.14	0.84	-5.356 **	8.7924 **
ELCEQ	253	0.14	1.23	0.1138	-0.10	0.77	-4.6912 **	7.0821 **
AUTOS	253	0.14	1.47	0.0952	0.14	0.75	-5.4472 **	7.2045 **
CARRY	253	0.10	1.15	0.0870	-0.32	0.88	-5.2106 **	13.103 ***
MINES	253	0.10	1.39	0.0719	0.23	0.74	-5.6247 **	8.6176 **
COAL	253	-0.22	2.10	-0.1048	-0.37	1.39	-5.1724 **	27.415 ***
OIL	253	-0.05	2.28	-0.0219	0.55	2.74	-5.3632 **	94.641 ***
UTIL	253	0.07	0.70	0.1000	-0.33	0.87	-7.0213 **	13.043 ***
TELCM	253	0.02	1.05	0.0190	0.11	1.11	-6.3514 **	14.281 ***
SERVS	253	0.10	0.95	0.1053	-0.30	1.97	-5.8318 **	46.46 ***
BUSEQ	253	0.15	1.03	0.1456	-0.33	1.09	-5.141 **	17.807 ***
PAPER	253	0.06	1.15	0.0522	0.03	0.43	-6.0542 **	2.1898
TRANS	253	0.05	1.26	0.0397	-0.04	0.46	-5.5597 **	2.5693
WHLSL	253	0.08	1.02	0.0784	-0.13	0.92	-5.7945 **	10.286 ***
RTAIL	253	0.05	1.25	0.0400	-0.08	0.72	-5.7551 **	6.1576 **
MEALS	253	0.06	0.82	0.0732	0.00	1.12	-6.1484 **	13.889 ***
FIN	253	0.10	0.82	0.1220	-0.49	1.48	-5.7066 **	34.256 ***
OTHER	253	0.08	0.95	0.0842	1.89	14.84	-6.9311 **	2519.2 ***

*Notes: Std, S/R, ADF and JB are the abbreviations of the standard deviation, Sharpe ratio, Augmented Dickey Fuller and Jarque-Bera tests, respectively. 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.*

4.4 Results

4.4.1 Clustering Analysis of Cryptocurrencies

First, we calculate the similarity of the distributions of daily log returns of two cryptocurrencies following Maasoumi & Racine (2002). 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). Block bootstrap is conducted via resampling with replacement from the pooled empirical distributions of X and Y under the Null of equality.

Tables 5a-5b present the distances across individual cryptocurrencies for the period from January 1, 2019 to January 1, 2021. The calculated S_ρ in a grey box indicates insignificance at the 95 percent level. Since there are too many significant values, we instead highlight the insignificant ones. From Table 5a-5b, we can clearly observe that many of the cryptocurrencies have different time distributions of daily returns, since many of the pair-wise distances S_ρ are significant. These results make it evident that there does not exist any market-wide convergence across cryptocurrency market.

Table 5a-5b provides detailed information about how one cryptocurrency is different from the remaining cryptocurrencies. Taking Bitcoin (BTC) as an example, in the first row of Table 5a, we first note that the distance S_ρs between Bitcoin and two other cryptocurrencies, Ethereum (ETH) and Litecoin (LTC), are significant at the 95 percent level, which means that the time distributions of the daily log returns of Bitcoin and two other cryptocurrencies were diverging. In particular, the distance between Bitcoin and Ethereum and Litecoin are 0.0185 and 0.0282, respectively. Similarly, Table 5a-5b also lists the other pair-wise distances between cryptocurrencies. We also plotted the heat map for entropy metrics S_ρ between cryptocurrencies in Figure 4, which illustrates which cryptocurrencies have large dissimilarities (blue color) versus those that appear to be fairly similar (red color).

As we mentioned above, although we reject that there exists a market-wide convergence, we find that there are still some distances between individual cryptocurrencies that are not significant and that are significant but very small (for example, the distance between Decred (DCR) and Tezos (XTZ) in Table 5b is 0.0094, although significant, but roughly zero). This finding makes us suspect that there might exist a club convergence. Clustering analysis allows us to find out whether or not this is indeed the case, and if so, how many groups or clubs do we have, and what are the members within each group?

Table 5a: Distances $D_{i,j}$ computed between each cryptocurrency series

	ETH	XRP	LTC	BCH	ADA	BNB	LINK	USDC	BSV	XLM	XMR	EOS	THETA	TRX	XEM
BTC	0.0185	0.0166	0.0282	0.0274	0.0406	0.0248	0.0717	0.4993	0.0312	0.0308	0.0303	0.0242	0.0847	0.0275	0.0351
ETH		0.0126	0.0079	0.0114	0.0128	0.0059	0.0323	0.4993	0.0119	0.0155	0.0066	0.0091	0.0413	0.0079	0.0164
XRP			0.0155	0.0156	0.0277	0.0161	0.0473	0.5000	0.0170	0.0164	0.0207	0.0153	0.0585	0.0169	0.0227
LTC				0.0079	0.0087	0.0080	0.0179	0.5000	0.0092	0.0076	0.0074	0.0129	0.0251	0.0073	0.0090
BCH					0.0157	0.0108	0.0247	0.5000	0.0104	0.0106	0.0140	0.0129	0.0322	0.0124	0.0098
ADA						0.0107	0.0129	0.5000	0.0150	0.0148	0.0109	0.0168	0.0166	0.0060	0.0113
BNB							0.0270	0.4993	0.0149	0.0129	0.0057	0.0147	0.0350	0.0083	0.0144
LINK								0.5000	0.0239	0.0219	0.0270	0.0338	0.0051	0.0202	0.0172
USDC									0.4993	0.5000	0.5000	0.4994	0.5000	0.5000	0.5000
BSV										0.0135	0.0180	0.0150	0.0316	0.0122	0.0099
XLM											0.0121	0.0168	0.0321	0.0125	0.0112
XMR												0.0148	0.0336	0.0099	0.0183
EOS													0.0418	0.0090	0.0175
THETA														0.0256	0.0240
TRX															0.0114
XEM															
XTZ															
CRO															
CEL															
VET															
DOGE															
SNX															
NEO															
MKR															
HT															
MIOTA															
DASH															
REV															
ZEC															
ETC															
DCR															

Table 5b: Distances $D_{i,j}$ computed between each cryptocurrency series (cont.)

	XTZ	CRO	CEL	VET	DOGE	SNX	NEO	MKR	HT	MIOTA	DASH	REV	ZEC	ETC	DCR
BTC	0.0500	0.0237	0.0981	0.0555	0.0150	0.1267	0.0364	0.0369	0.0066	0.0314	0.0217	0.0362	0.0394	0.0182	0.0301
ETH	0.0190	0.0138	0.0495	0.0225	0.0233	0.0723	0.0139	0.0130	0.0123	0.0067	0.0127	0.0538	0.0132	0.0095	0.0065
XRP	0.0297	0.0171	0.0682	0.0356	0.0167	0.0939	0.0232	0.0201	0.0143	0.0146	0.0142	0.0427	0.0225	0.0123	0.0188
LTC	0.0086	0.0106	0.0325	0.0123	0.0300	0.0516	0.0058	0.0083	0.0232	0.0062	0.0103	0.0596	0.0061	0.0104	0.0049
BCH	0.0156	0.0121	0.0397	0.0186	0.0268	0.0601	0.0098	0.0122	0.0219	0.0096	0.0128	0.0561	0.0098	0.0111	0.0099
ADA	0.0080	0.0158	0.0236	0.0074	0.0408	0.0386	0.0079	0.0091	0.0319	0.0092	0.0164	0.0709	0.0068	0.0160	0.0044
BNB	0.0155	0.0133	0.0441	0.0186	0.0284	0.0677	0.0110	0.0094	0.0202	0.0061	0.0105	0.0628	0.0092	0.0122	0.0046
LINK	0.0067	0.0270	0.0091	0.0058	0.0682	0.0176	0.0124	0.0153	0.0616	0.0213	0.0321	0.0969	0.0150	0.0323	0.0192
USDC	0.5000	0.4980	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.4993	0.5000	0.5000	0.4993	0.5000	0.5000	0.5000
BSV	0.0139	0.0116	0.0355	0.0181	0.0296	0.0532	0.0124	0.0144	0.0244	0.0118	0.0140	0.0551	0.0112	0.0130	0.0117
XLM	0.0140	0.0118	0.0403	0.0167	0.0297	0.0582	0.0089	0.0083	0.0254	0.0118	0.0135	0.0609	0.0101	0.0136	0.0101
XMR	0.0170	0.0158	0.0457	0.0193	0.0362	0.0675	0.0096	0.0103	0.0251	0.0063	0.0170	0.0707	0.0103	0.0135	0.0055
EOS	0.0205	0.0145	0.0481	0.0218	0.0196	0.0674	0.0164	0.0184	0.0158	0.0130	0.0185	0.0453	0.0167	0.0083	0.0109
THETA	0.0102	0.0347	0.0058	0.0080	0.0820	0.0125	0.0193	0.0216	0.0737	0.0267	0.0421	0.1092	0.0193	0.0424	0.0253
TRX	0.0108	0.0128	0.0317	0.0112	0.0277	0.0502	0.0089	0.0098	0.0212	0.0073	0.0119	0.0564	0.0093	0.0086	0.0041
XEM	0.0121	0.0136	0.0305	0.0150	0.0327	0.0476	0.0093	0.0097	0.0285	0.0129	0.0106	0.0606	0.0095	0.0121	0.0094
XTZ		0.0151	0.0150	0.0033	0.0472	0.0286	0.0077	0.0109	0.0408	0.0113	0.0194	0.0732	0.0081	0.0203	0.0094
CRO			0.0420	0.0189	0.0250	0.0602	0.0115	0.0166	0.0191	0.0138	0.0128	0.0480	0.0141	0.0121	0.0129
CEL				0.0120	0.0903	0.0065	0.0258	0.0326	0.0838	0.0354	0.0519	0.1175	0.0273	0.0497	0.0328
VET					0.0529	0.0235	0.0095	0.0133	0.0455	0.0137	0.0251	0.0797	0.0098	0.0240	0.0115
DOGE						0.1174	0.0380	0.0400	0.0104	0.0328	0.0257	0.0393	0.0370	0.0175	0.0310
SNX							0.0436	0.0498	0.1105	0.0562	0.0721	0.1406	0.0450	0.0708	0.0525
NEO								0.0071	0.0317	0.0079	0.0138	0.0697	0.0076	0.0139	0.0068
MKR									0.0312	0.0068	0.0124	0.0699	0.0087	0.0136	0.0072
HT										0.0250	0.0197	0.0373	0.0310	0.0152	0.0224
MIOTA											0.0120	0.0610	0.0070	0.0119	0.0055
DASH												0.0496	0.0143	0.0110	0.0117
REV													0.0676	0.0455	0.0649
ZEC														0.0157	0.0067
ETC															0.0099
DCR															

Notes: Figures highlighted are insignificant at $p <= 0.05$ *level.*

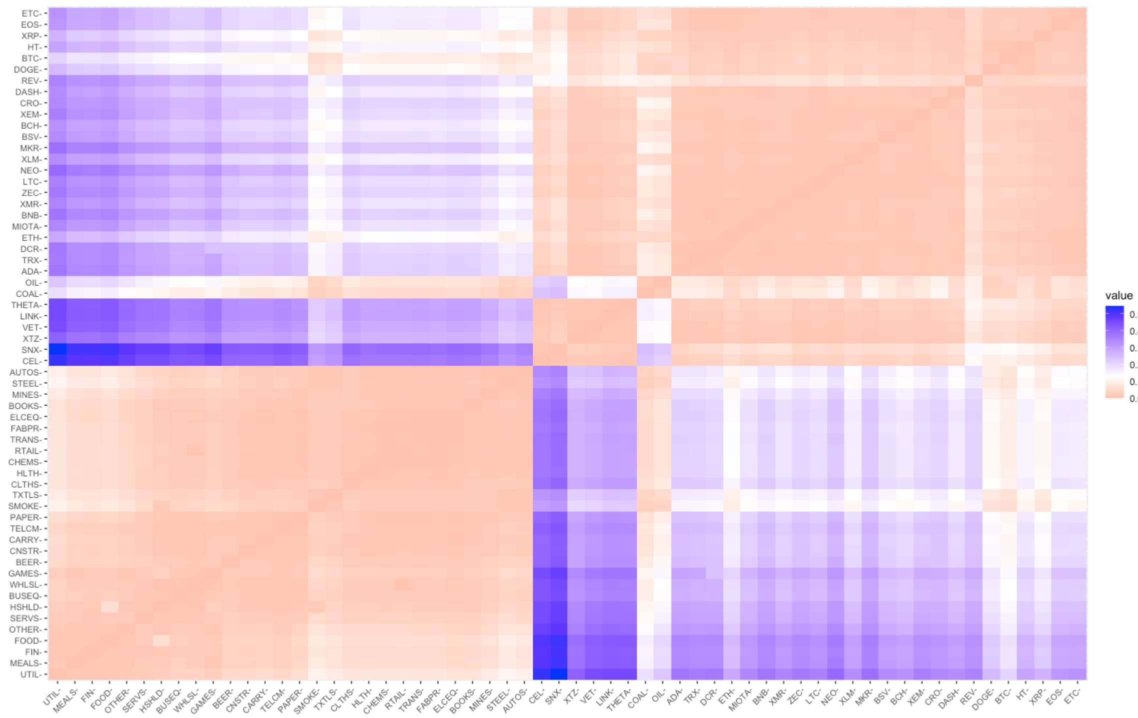


Figure 4.4: Heat map for entropy metrics between cryptocurrencies and commodities

The results obtained from hierarchical clustering analysis using three different linkage (complete, single and average) are summarized in Figures 5. The figure 5 demonstrates how our hierarchical clustering method proceeds, combining the clusters beginning with the case where each cryptocurrencies is its own cluster to the case where all the cryptocurrencies are in the same cluster. To determine the optimal number of clusters, we look at the largest difference of heights in the dendrogram. The optimal number of clusters is 4 for complete and average linkage and the optimal number of clusters is 2 for single linkage in our case. Using complete linkage method, we can cluster the cryptocurrencies into four clubs, and the membership of each club is summarized in Table 6.

Table 4.6: Clustering Analysis Results for Cryptocurrencies (complete linkage)

1	CEL	SNX	XTZ	VET	LINK	THETA
2	REV					
3	BSV	BCH	XEM	CRO	DASH	XLM
	MKR	ETH	MIOTA	BNB	XMR	ADA
	TRX	DCR	ZEC	LTC	NEO	
4	DOGE	BTC	HT	XRP	EOS	ETC

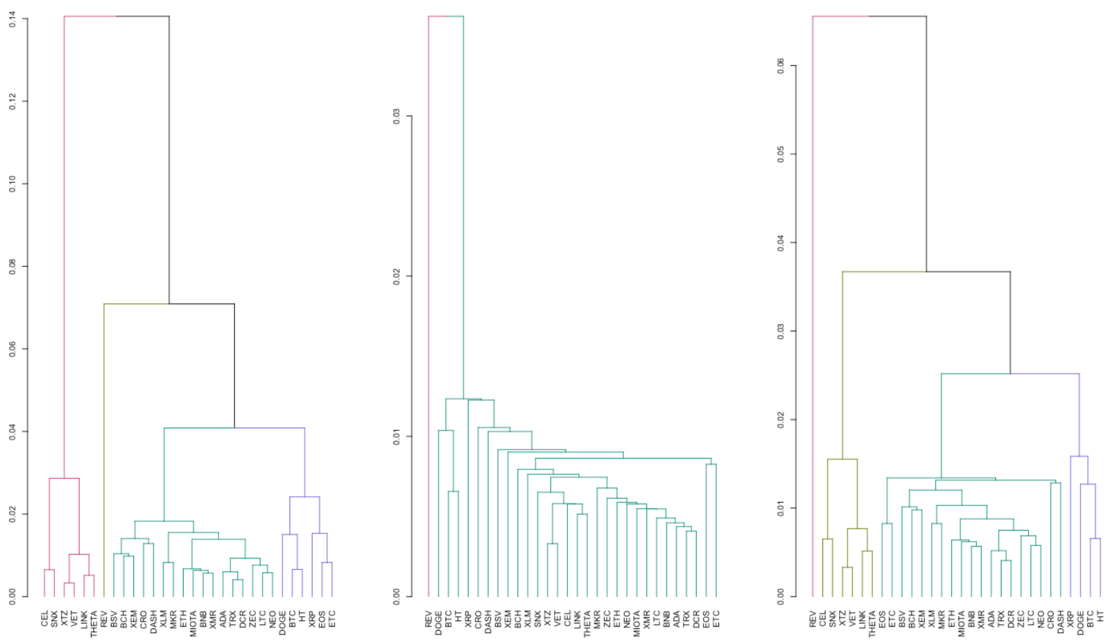


Figure 4.5: Dendrogram with complete (left), single (middle) and average (right) linkage for clusters consist of cryptocurrencies

- **Cluster 1**: This cluster allocates 6 cryptocurrencies with very high average returns, high volatility as well as high Sharpe Ratios. This cluster includes: Celsius (CEL), Synthetix (SNX), Tezos (XTZ), VeChain (VET), Chainlink (LINK) and THETA (TEHTA). The average daily log-returns is 0.48%, the average volatility is 6.84, and the average Sharpe Ratio is 0.07 within this cluster. We can also observe that the cryptocurrencies in this cluster have low skewness and kurtosis. We also notice that most of the cryptocurrencies in this cluster have low volume.

- **Cluster 2**: This cluster contains only one cryptocurrency, Revain (REV), which have negative return (also negative Sharpe Ratio) and high skewness and kurtosis. The average daily log-return of REV is -0.40%, the average volatility of REV is 5.40, and the average Sharpe Ratio of REV is -0.07. REV was introduced in the market in 2017 as an Ethereum-based token that interfaces with the Revain online business review platform. The organisation utilizes the Ethereum blockchain to record reviews approved by both the system and the business owner.

- **Cluster 3**: This cluster contains cryptocurrencies with low returns and low volatility, as well as mild skewness and kurtosis. This cluster includes: Bitcoin SV (BSV), Bitcoin Cash (BCH), NEM (XEM), Crypto.com Coin (CRO), Dash (DASH), Stellar (XLM), Maker (MKR), Ethereum (ETH), IOTA (MIOTA), Binance Coin (BNB), Monero (XMR), Cardano (ADA), TRON (TRX), Decred (DCR), Zcash (ZEC), Litecoin (LTC) and Neo (NEO). The average daily log-returns is 0.11%, the average volatility is 5.32, and the average Sharpe Ratio is 0.02 within this cluster. One of the cryptocurrencies with highest market capitalization, Ethereum (ETH), is also included in this cluster. The similarities among the cryptocurrencies within the cluster can provide possible explanations: (1) Geographic similarity. For example, we noticed that both Cardano (ADA) and Neo (NEO) have projects in Asia and most of their investors are located in Asia, which leads them to be classified in the same cluster. (2) Function. Dash (DASH) and Monero (XMR) are in the same cluster for the potential reason that these two cryptocurrencies emphasized on privacy and security. (3) "Coinbase effect". We found that during a certain listing time, the cryptocurrencies under the same market conditions can show similar trends. For example, Coinbase announced that Zcash (ZEC) was listed in November 2018, and after that, two other cryptocurrencies Stellar (XLM) and Maker (MKR) were also listed on Coinbase within few months. This causes these three cryptocurrencies to be classified into the same cluster.

- **Cluster 4**: This cluster contains cryptocurrencies with low returns and low volatility. This cluster includes: Dogecoin (DOGE), Bitcoin (BTC), Huobi Token (HT), XRP (XRP), EOS (EOS) and Ethereum Classic (ETC). The average daily log-returns is 0.10%, the average volatility is 4.59, and the average Sharpe Ratio is 0.02 within this cluster. This cluster contains the cryptocurrency with highest market capitalization, Bitcoin (BTC). By observing this cluster, our assumption on geographic similarity is also strengthened. Dogecoin (DOGE) and XRP (XRP) belonged to the same cluster, and both of them are digital currencies and most of their investors are based in US.

4.4.2 Clustering Analysis of Cryptocurrencies and Commodities

In Maasoumi & Wu (2021), we analyzed the density similarity between Bitcoin and select asset classes, and found that the top three industries which have similar densities with Bitcoin are Coal, Steel and Mines. This novel finding inspired us to construct a portfolio that can best simulate the performance of Bitcoin return using these three industries of stocks. In Maasoumi & Wu (2021), we calculated the entropy measures between Bitcoin and select asset classes before and after COVID-19 broke out. Before COVID-19 broke out, 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. After COVID-19 broke out, the density of Bitcoin daily return has smallest distance with the density of Business Supplies and Shipping Containers, Utilities, Tobacco Products and Restaurants, Hotels, Motels industries daily returns. Comparing S_ρ before and after the broke out of COVID-19, we conclude that the values of S_ρ decrease generally in all cases.

The previous findings inspired us to implement hierarchical clustering techniques to examine if the clusters of cryptocurrencies we derived in Section 4.1 include other asset classes, such as commodities.

The data for commodities comes from Kenneth R. French 30 Industry Portfolios. We use the daily average value weighted returns for 30 industry portfolios data. The 30 commodities that included in our analysis are listed in Table 2.

Table 7a-7b listed the S_ρ between cryptocurrencies and commodities, which provides detailed information about how one cryptocurrency is different from one commodity asset. All the S_ρs we calculated are statistically significant at the 95 percent level, which means that the time distributions of the daily log returns of cryptocurrencies and commodities were diverging. Table 7a-7b show the distance between each pair-wise cryptocurrency and commodity. For example, in row 1 of Table 7a, the S_ρ between Bitcoin (BTC) and Food Products (Food) is 0.1837 and significant at the 95 percent level. This indicates that the time distributions of the daily log returns of Bitcoin and Food Products were diverging, and distance between their densities is

0.1837.

Table 7a: Distances $D_{i,j}$ computed between each cryptocurrency and commodity series

	FOOD	BEER	SMOKE	GAMES	BOOKS	HSHLD	CLTHS	HLTH	CHEMS	TXTLS	CNSTR	STEEL	FABPR	ELCEQ	AUTOS
BTC	0.1837	0.1153	0.0559	0.1413	0.0944	0.1419	0.1021	0.0941	0.0897	0.0645	0.1111	0.0651	0.0875	0.0906	0.0742
ETH	0.2576	0.1698	0.0885	0.2052	0.1378	0.2044	0.1501	0.1328	0.1337	0.0995	0.1595	0.0977	0.1337	0.1405	0.1085
XRP	0.2298	0.1435	0.0720	0.1792	0.1144	0.1785	0.1243	0.1137	0.1118	0.0844	0.1402	0.0807	0.1107	0.1164	0.0923
LTC	0.3287	0.2318	0.1321	0.2735	0.1928	0.2730	0.2060	0.1894	0.1893	0.1495	0.2222	0.1437	0.1865	0.1951	0.1562
BCH	0.2993	0.2052	0.1161	0.2455	0.1704	0.2459	0.1843	0.1689	0.1680	0.1285	0.1982	0.1267	0.1669	0.1742	0.1398
ADA	0.3483	0.2533	0.1496	0.2947	0.2153	0.2940	0.2266	0.2109	0.2116	0.1670	0.2440	0.1655	0.2091	0.2174	0.1794
BNB	0.3473	0.2538	0.1456	0.2951	0.2150	0.2952	0.2271	0.2100	0.2113	0.1615	0.2421	0.1630	0.2088	0.2175	0.1766
LINK	0.4278	0.3330	0.2235	0.3785	0.2983	0.3782	0.3085	0.2945	0.2957	0.2426	0.3283	0.2460	0.2929	0.3016	0.2622
BSV	0.3076	0.2185	0.1313	0.2584	0.1868	0.2570	0.1945	0.1845	0.1839	0.1486	0.2152	0.1461	0.1809	0.1881	0.1587
XLM	0.3038	0.2027	0.1119	0.2460	0.1634	0.2454	0.1775	0.1650	0.1623	0.1310	0.1995	0.1193	0.1603	0.1682	0.1344
XMR	0.3206	0.2216	0.1248	0.2632	0.1848	0.2619	0.1947	0.1794	0.1789	0.1427	0.2121	0.1353	0.1771	0.1852	0.1487
EOS	0.2907	0.2018	0.1205	0.2371	0.1698	0.2369	0.1829	0.1654	0.1654	0.1314	0.1919	0.1289	0.1652	0.1715	0.1400
THETA	0.4272	0.3312	0.2214	0.3757	0.2979	0.3757	0.3071	0.2909	0.2921	0.2391	0.3227	0.2423	0.2892	0.2973	0.2572
TRX	0.3402	0.2465	0.1484	0.2891	0.2112	0.2886	0.2244	0.2094	0.2100	0.1639	0.2416	0.1658	0.2086	0.2170	0.1802
XEM	0.3292	0.2370	0.1355	0.2752	0.1969	0.2742	0.2069	0.1942	0.1932	0.1540	0.2273	0.1494	0.1900	0.1978	0.1634
XTZ	0.3876	0.2947	0.2005	0.3396	0.2662	0.3393	0.2767	0.2622	0.2636	0.2148	0.2926	0.2199	0.2624	0.2703	0.2345
CRO	0.3191	0.2324	0.1524	0.2743	0.2068	0.2735	0.2157	0.2043	0.2055	0.1661	0.2326	0.1689	0.2050	0.2125	0.1820
CEL	0.4985	0.4041	0.3017	0.4521	0.3737	0.4523	0.3850	0.3714	0.3722	0.3215	0.4041	0.3230	0.3694	0.3780	0.3388
VET	0.4193	0.3215	0.2055	0.3661	0.2814	0.3660	0.2928	0.2775	0.2780	0.2263	0.3127	0.2257	0.2747	0.2837	0.2419
DOGE	0.2295	0.1476	0.0690	0.1809	0.1133	0.1809	0.1286	0.1145	0.1139	0.0799	0.1407	0.0791	0.1135	0.1201	0.0910
SNX	0.5187	0.4268	0.3149	0.4729	0.3945	0.4732	0.4050	0.3892	0.3916	0.3324	0.4217	0.3398	0.3889	0.3978	0.3556
NEO	0.3672	0.2755	0.1583	0.3150	0.2368	0.3131	0.2386	0.2264	0.2283	0.1801	0.2608	0.1800	0.2250	0.2335	0.1955
MKR	0.3590	0.2654	0.1633	0.3069	0.2308	0.3067	0.2413	0.2254	0.2262	0.1798	0.2563	0.1805	0.2240	0.2319	0.1943
HT	0.2677	0.1836	0.1023	0.2184	0.1536	0.2173	0.1633	0.1480	0.1491	0.1135	0.1738	0.1146	0.1479	0.1545	0.1242
MIOTA	0.3176	0.2208	0.1249	0.2627	0.1844	0.2626	0.1960	0.1807	0.1809	0.1404	0.2131	0.1373	0.1798	0.1878	0.1513
DASH	0.3053	0.2114	0.1114	0.2515	0.1700	0.2506	0.1836	0.1698	0.1692	0.1294	0.2039	0.1249	0.1668	0.1754	0.1392
REV	0.3279	0.2430	0.1505	0.2809	0.2097	0.2801	0.2186	0.2074	0.2079	0.1683	0.2374	0.1682	0.2055	0.2128	0.1818
ZEC	0.3417	0.2418	0.1419	0.2878	0.2022	0.2880	0.2194	0.2030	0.2034	0.1577	0.2384	0.1559	0.2022	0.2115	0.1711
ETC	0.2933	0.2021	0.1092	0.2394	0.1675	0.2382	0.1751	0.1621	0.1616	0.1259	0.1921	0.1216	0.1592	0.1658	0.1347
DCR	0.3414	0.2416	0.1418	0.2394	0.2026	0.2852	0.2160	0.2016	0.2006	0.1609	0.2362	0.1542	0.1981	0.2064	0.1692

Table 7b: Distances $D_{i,j}$ computed between each cryptocurrency and commodity series (cont.)

	CARRY	MINES	COAL	OIL	UTIL	TELCM	SERVS	BUSEQ	PAPER	TRANS	WHLSL	RTAIL	MEALS	FIN	OTHER
BTC	0.1085	0.0843	0.0376	0.0397	0.2071	0.1168	0.1530	0.1312	0.1091	0.0921	0.1314	0.0908	0.1783	0.1782	0.1631
ETH	0.1624	0.1196	0.0481	0.0429	0.2826	0.1676	0.2110	0.1868	0.1547	0.1355	0.1886	0.1345	0.2506	0.2482	0.2289
XRP	0.1360	0.1038	0.0380	0.0368	0.2660	0.1434	0.1858	0.1614	0.1320	0.1127	0.1627	0.1115	0.2243	0.2217	0.2033
LTC	0.2208	0.1719	0.0808	0.0643	0.3584	0.2313	0.2789	0.2501	0.2164	0.1909	0.2537	0.1899	0.3235	0.3195	0.2998
BCH	0.1991	0.1536	0.0694	0.0590	0.3258	0.2048	0.2521	0.2266	0.1922	0.1699	0.2282	0.1678	0.2927	0.2911	0.2686
ADA	0.2424	0.1953	0.0919	0.0835	0.3774	0.2539	0.2996	0.2714	0.2389	0.2136	0.2752	0.2122	0.3435	0.3392	0.3179
BNB	0.2441	0.1934	0.0939	0.0770	0.3762	0.2533	0.2993	0.2726	0.2397	0.2142	0.2762	0.2118	0.3429	0.3395	0.3180
LINK	0.3269	0.2797	0.1581	0.1444	0.4556	0.3381	0.3815	0.3556	0.3246	0.2985	0.3596	0.2958	0.4245	0.4201	0.3979
BSV	0.2091	0.1737	0.0820	0.0780	0.3354	0.2211	0.2621	0.2358	0.2082	0.1854	0.2397	0.1838	0.3044	0.2992	0.2810
XLM	0.1925	0.1493	0.0563	0.0466	0.3331	0.2037	0.2534	0.2241	0.1884	0.1628	0.2262	0.1619	0.2976	0.2939	0.2757
XMR	0.2109	0.1641	0.0690	0.0592	0.3493	0.2220	0.2694	0.2397	0.2055	0.1812	0.2435	0.1798	0.3137	0.3100	0.2908
EOS	0.1940	0.1510	0.0723	0.0711	0.3162	0.2000	0.2452	0.2195	0.1865	0.1667	0.2207	0.1663	0.2831	0.2817	0.2607
THETA	0.3245	0.2746	0.1575	0.1410	0.4557	0.3356	0.3808	0.3529	0.3211	0.2954	0.3569	0.2929	0.4225	0.4194	0.3958
TRX	0.2406	0.1949	0.0929	0.0823	0.3667	0.2484	0.2924	0.2681	0.2354	0.2120	0.2710	0.2099	0.3355	0.3315	0.3095
XEM	0.2222	0.1795	0.0807	0.0763	0.3588	0.2352	0.2808	0.2516	0.2201	0.1949	0.2555	0.1934	0.3246	0.3199	0.2998
XTZ	0.2935	0.2494	0.1421	0.1323	0.4128	0.3011	0.3430	0.3198	0.2890	0.2664	0.3225	0.2637	0.3832	0.3800	0.3576
CRO	0.2323	0.1954	0.1088	0.1005	0.3420	0.2381	0.2758	0.2558	0.2281	0.2083	0.2586	0.2053	0.3153	0.3114	0.2933
CEL	0.4028	0.3558	0.2339	0.2136	0.5248	0.4130	0.4549	0.4306	0.4003	0.3748	0.4342	0.3722	0.4954	0.4916	0.4712
VET	0.3112	0.2602	0.1380	0.1210	0.4492	0.3235	0.3706	0.3419	0.3085	0.2806	0.3460	0.2783	0.4152	0.4111	0.3891
DOGE	0.1404	0.1019	0.0358	0.0300	0.2533	0.1442	0.1858	0.1647	0.1341	0.1152	0.1659	0.1135	0.2241	0.2215	0.2028
SNX	0.4237	0.3732	0.2467	0.2273	0.5450	0.4336	0.4751	0.4510	0.4208	0.3950	0.4551	0.3920	0.5158	0.5121	0.4888
NEO	0.2599	0.2152	0.1064	0.0885	0.3967	0.2748	0.3177	0.2882	0.2593	0.2331	0.2946	0.2293	0.3634	0.3578	0.3403
MKR	0.2575	0.2101	0.1075	0.0960	0.3869	0.2670	0.3119	0.2851	0.2531	0.2289	0.2885	0.2268	0.3539	0.3507	0.3289
HT	0.1757	0.1361	0.0635	0.0601	0.2929	0.1823	0.2227	0.1986	0.1701	0.1511	0.2017	0.1499	0.2622	0.2589	0.2389
MIOTA	0.2127	0.1666	0.0690	0.0607	0.3454	0.2212	0.2686	0.2417	0.2073	0.1831	0.2441	0.1811	0.3116	0.3088	0.2885
DASH	0.1991	0.1545	0.0605	0.0529	0.3341	0.2098	0.2557	0.2286	0.1956	0.1705	0.2319	0.1690	0.3008	0.2959	0.2764
REV	0.2339	0.1963	0.0969	0.0942	0.3546	0.2438	0.2840	0.2600	0.2322	0.2097	0.2630	0.2078	0.3251	0.3203	0.3019
ZEC	0.2359	0.1864	0.0828	0.0693	0.3698	0.2440	0.2919	0.2666	0.2305	0.2046	0.2686	0.2031	0.3370	0.3332	0.3117
ETC	0.1895	0.1494	0.0621	0.0556	0.3221	0.2013	0.2461	0.2170	0.1867	0.1636	0.2204	0.1623	0.2878	0.2840	0.2644
DCR	0.2317	0.1851	0.0820	0.0719	0.3712	0.2433	0.2918	0.2627	0.2281	0.2018	0.2656	0.2007	0.3362	0.3322	0.3119

Notes: Figures highlighted are insignificant at $p <= 0.05$ *level.*

Next, we explore whether cryptocurrencies and commodities could form converging clubs using hierarchical clustering analysis. In other words, we would like to analyze whether the most similar commodity assets that we have identified with Bitcoin will belong to the same "clusters" or clubs. We conduct the hierarchical clustering analysis using three different linkage (complete, single and average) are summarized in Figures 6. The figure 6 demonstrates how our hierarchical clustering method proceeds, combining the clusters beginning with the case where each cryptocurrency /

commodity assets is its own cluster to the case where all the assets are in the same cluster. The optimal number of clusters is 8 for complete linkage, 4 for single linkage and 6 for average linkage in our case. Using complete linkage method, we can cluster the cryptocurrencies into 8 clubs, and the components of each club are summarized in Table 8.

From Table 8, we noticed that Cluster 1, 2, 3 and 5 only contain commodities, while Cluster 4, 6, 7 and 8 only contain cryptocurrencies, and the classification of cryptocurrencies are exactly the same as when we cluster cryptocurrencies only (see Table 6). Camparing with commodities, cryptocurrencies returns exhibit high variability and excess kurtosis. Hence, cryptocurrencies and commodities converge to separated clusters intuitively. In addition, we noticed that Cluster 5, which contains commodities Coal (COAL) and Petroleum and Natural Gas (OIL) have smaller distance to cryptocurrencies according to our cluster proximity using entropy metric S_ρ. This result is consistent with our previous analysis in Maasoumi & Wu (2019), which discovered top density similarity between Bitcoin and Coal, Steel and Mines industries. This finding will be useful for diversification among different asset classes.

Table 4.8: Clustering Analysis Results for Cryptocurrencies and Commodities (complete linkage)

1	UTIL	MEALS	FIN	FOOD	OTHER	
2	CLTHS	HLTH	CHEMS	RTAIL	TRANS	FABPR
	ELCEQ	SMOKE	TXTLS	BOOKS	MINES	STEEL
	AUTOS					
3	SERVS	HSHLD	BUSEQ	WHLSL	GAMES	BEER
	CNSTR	CARRY	TELCM	PAPER		
4	CEL	SNX	XTZ	VET	LINK	THETA
5	COAL	OIL				
6	REV					
7	BSV	BCH	XEM	CRO	DASH	XLM
	MKR	ETH	MIOTA	BNB	XMR	ADA
	TRX	DCR	ZEC	LTC	NEO	
8	DOGE	BTC	HT	XRP	EOS	ETC

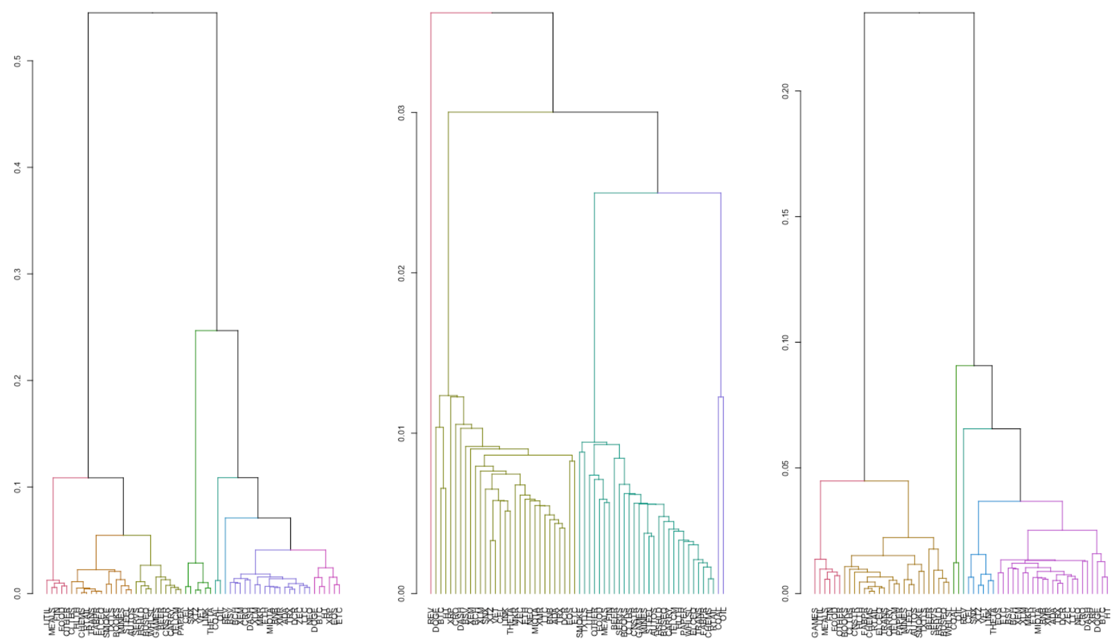


Figure 4.6: Dendrogram with complete (left), single (middle) and average (right) linkage for clusters consist of cryptocurrencies and commodities

4.4.3 Comparing with K-means Clustering Results

In the previous section, we have shown that we can get 4 clusters of cryptocurrencies with entropy metric $\S_\rho$ as measurement of similarity using Agglomerative Hierarchical Clustering with complete linkage. Similarly, we can get 8 clusters if we consider both cryptocurrencies and commodities simultaneously.

Both Hierarchical Clustering and K-means Clustering algorithms look for similarities among data and both use the same approaches to decide the number of clusters. However, they are slightly different in dealing with different data set: K-means is method of cluster analysis using a pre-specified number of clusters, which requires advance knowledge of "K"; While Hierarchical clustering seeks to build a hierarchy of clusters without having fixed number of cluster. The main differences between the two clustering methods are: (1) K-means Clustering uses a pre-specified number of clusters, the method assigns records to each cluster to find the mutually exclusive cluster of spherical shape based on distance; Hierarchical Clustering can be either divisive or agglomerative. (2) K-means Clustering needs advance knowledge of K (number of clusters one want to divide your data); In Hierarchical Clustering one can stop at any number of clusters, one find appropriate by interpreting the dendrogram. (3) In K-means Clustering, one can use median or mean as a cluster centre to represent each cluster; Agglomerative Hierarchical Clustering begins with "n" clusters and sequentially combine similar clusters until only one cluster is obtained. (4) In K-means Clustering, methods used are normally less computationally intensive and are suited with very large data sets; In Hierarchical Clustering, divisive methods work in the opposite direction, beginning with one cluster that includes all the records and Hierarchical methods are especially useful when the target is to arrange the clusters into a natural hierarchy. (5) In K-means Clustering, since one start with random choice of clusters, the results produced by running the algorithm many times may differ; In Hierarchical Clustering, results are reproducible in Hierarchical clustering. (6) K-means Clustering is found to work well when the structure of the clusters is hyper spherical (like circle in 2D, sphere in 3D); Hierarchical Clustering does not work as well as K-means when the shape of the clusters is hyper spherical. The main advantages for K-means Clustering are: Convergence is guaranteed; Specialized to clusters of different sizes and shapes. While the main disadvantages for K-means Clustering are: The "K" is difficult to predict; Didn't work well with global cluster. The main advantages for Hierarchical Clustering are: Ease of handling of any forms of similarity or distance; Applicability to any attributes types. While the main disadvantages for Hierarchical Clustering are: Hierarchical clustering requires the computation and storage of an $n \times n$ distance matrix; For very large data sets, this can be expensive and slow. Therefore, in this section, we aim to compare our results using Agglomerative

Hierarchical Clustering in the previous section and using K-means Clustering, which can justify and strengthen our conclusions.

Firstly, we conduct K-means cluster analysis for portfolio only containing cryptocurrencies. We first explore how many clusters can be distinguished in data set. For this purpose, we use the elbow method for all defined feature sets. We found that there is a sharp elbow at $k = 2$, therefore, we determine the optimal number of clusters is 2. Table 9 and Figure 8 shows the K-means cluster analysis result for portfolio only containing cryptocurrencies.

Table 4.9: K-means Clustering Analysis Results for Cryptocurrencies

1	CEL	VET	DOGE	SNX	NEO	MKR
	HT	MIOTA	DASH	REV	ZEC	ETC
	DCR					
2	BTC	ETH	XRP	LTC	BCH	
	ADA	BNB	LINK	BSV	XLM	XMR
	EOS	THETA	TRX	XEM	XTZ	CRO

Next, we conduct K-means cluster analysis for portfolio containing both cryptocurrencies and commodities. Implementing elbow method, we found that there is a sharp elbow at $k = 9$, thus the optimal number of clusters should be 9. The K-means cluster analysis results are shown in Table 10 and Figure 10. Comparing cluster analysis results using K-means and our Agglomerative Hierarchical Clustering method with entropy metrics as distance measurement, we found that the K-means method divides the cryptocurrencies and commodities into separate clusters, specifically, all cryptocurrencies are classified into cluster 1, 2, 3, 4, 5, 6 and 8, while all commodities are classified into cluster 7 and 9. While with the Agglomerative Hierarchical Clustering method as we discuss in the previous sections, cryptocurrencies and commodities can be mixed up into same “clubs” if we tolerate fewer clusters. For example, if we tolerate only 3 clusters, then commodities Coal and Oil will be in the same cluster with cryptocurrencies.

Table 4.10: K-means Clustering Analysis Results for Cryptocurrencies and Commodities

1	LINK					
2	BSV					
3	BTC	ETH	XRP	LTC	BCH	ADA
	BNB	XLM	XMR	EOS	THETA	TRX
	XEM	XTZ				
4	SNX					
5	VET	DOGE	NEO	MKR	HT	MIOTA
	DASH	REV	ZEC	ETC	DCR	
6	CEL					
7	CHEMS	STEEL	FABPR	AUTOS	COAL	OIL
	TRANS					
8	CRO					
9	FOOD	BEER	SMOKE	GAMES	BOOKS	HSHLD
	CLTHS	HLTH	TXTLS	CNSTR	ELCEQ	CARRY
	MINES	UTIL	TELCM	SERVS	BUSEQ	PAPER
	WHLSL	RTAIL	MEALS	FIN	OTHER	

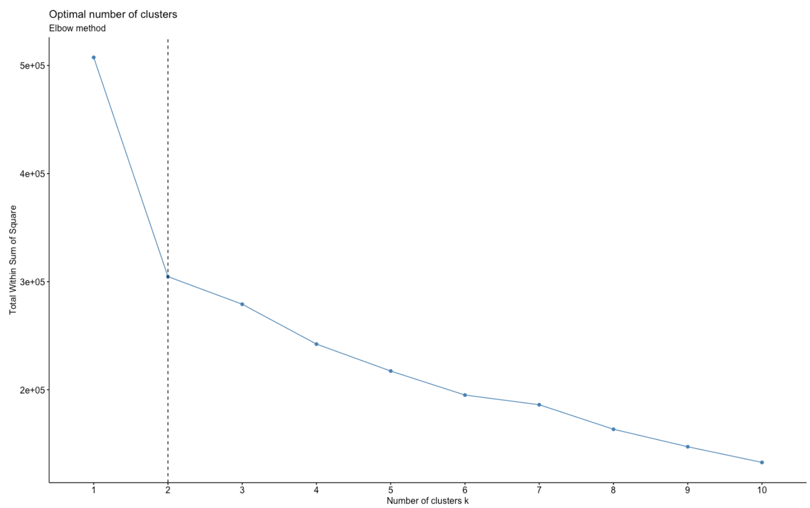


Figure 4.7: Elbow method to determine the optimal number of clusters consist of cryptocurrencies

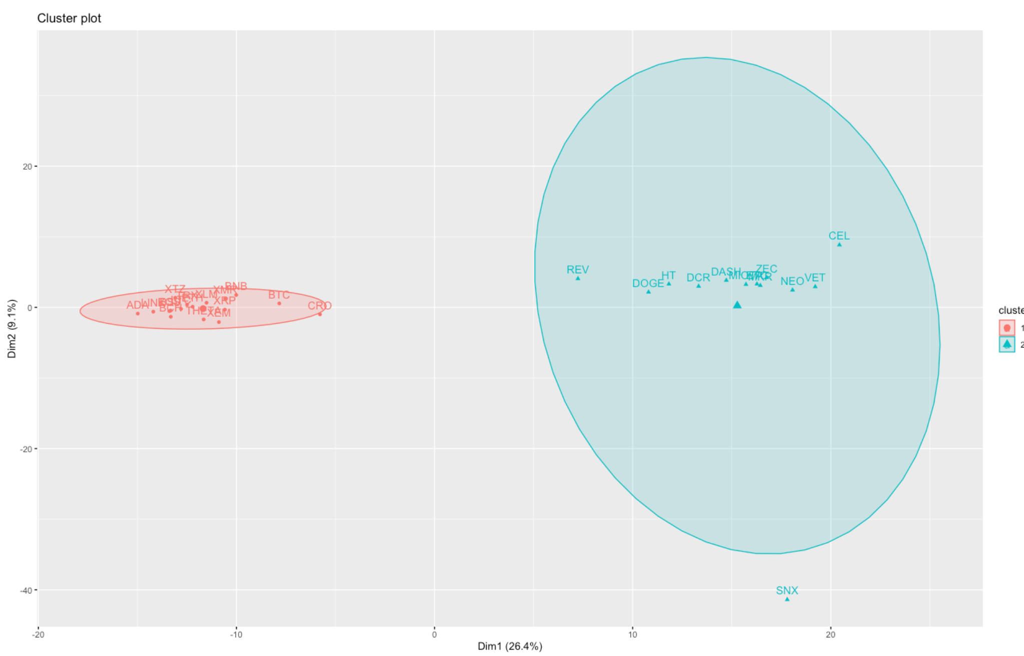


Figure 4.8: K-means cluster plot consist of cryptocurrencies

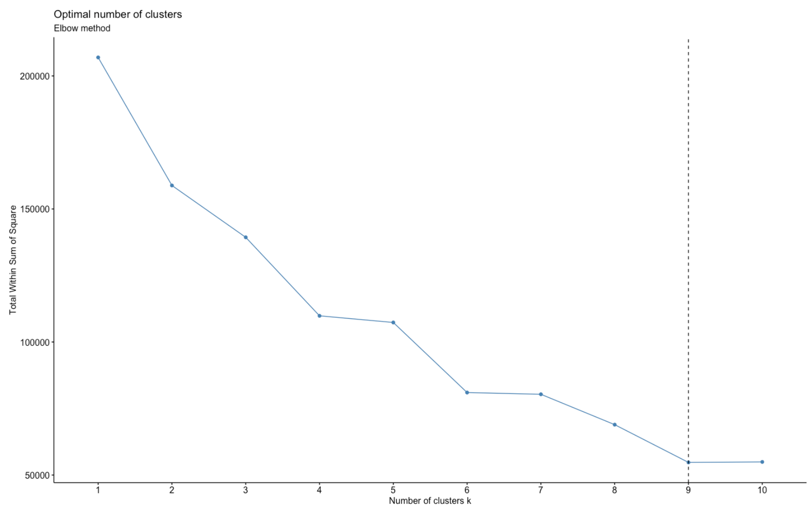


Figure 4.9: Elbow method to determine the optimal number of clusters consist of cryptocurrencies and commodities

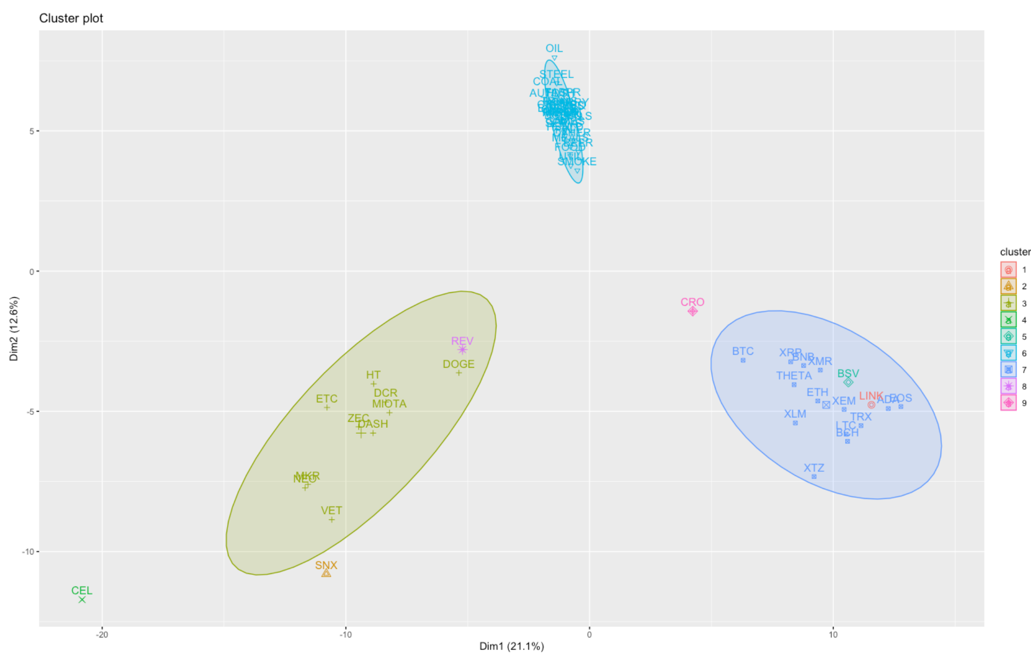


Figure 4.10: K-means cluster plot consist of cryptocurrencies and commodities

4.5 Conclusions

In this work, we analyzed the cryptocurrency market that consists cryptocurrencies with top 50 highest market capitalization, from January 1, 2019 to January 1, 2021, with a novel method that implementing Agglomerative Hierarchical Clustering (AHC) and K-means Clustering techniques based on the measurement of similarity entropy metrics S_ρ. We found that cryptocurrencies converge into four different clusters using AHC while converge into two clusters using K-means Clustering. We discover that cryptocurrencies share similar geographic locations tend to converge to same "clubs". For instance, Cardano (ADA) and Neo (NEO) belong to the same cluster since both of them have projects in Asia. We also notice that cryptocurrencies converge to same cluster when they share similar functions. For example, Dash (DASH) and Monero (XMR) are in the same cluster and they both focus on privacy and security. We also observe another potential explanation to our results called "Coinbase effect". In addition, we also examined if these clusters include other asset classes, such as commodities. We find the cluster that contains commodities Coal (COAL) and Petroleum and Natural Gas (OIL) have smaller distance to cryptocurrencies using entropy metric as cluster proximity, which is consistent with our previous work.

To conclude, we believe that the methodology that we used provides a consistent and descriptive tool supported by modern clustering techniques that may be useful for investors that need to understand the cryptocurrency market, as it depicts the entire distributions instead of only moments and identify the main trends in a descriptive manner.

For further investigations, the associations of some of the key financial ratios and cluster associations could play an important role enhancing the performance of the algorithms for the asset selection and diversification of portfolios (Brauneis and Mestel, 2019; Liu, 2019; Platanakis et al., 2018)) or improving the forecasting performance (Mallikarjuna and Rao, 2019) to tackle the difficulty of a new market.

Bibliography

[1] Brauneis, Alexander, and Roland Mestel. "Cryptocurrency-portfolios in a mean-variance framework." Finance Research Letters 28 (2019): 259-264.

[2] Burniske, Chris, and Jack Tatar. Cryptoassets: The innovative investor's guide to bitcoin and beyond. New York: McGraw-Hill Education, 2018.

[3] Corbet, Shaen, Andrew Meegan, Charles Larkin, Brian Lucey, and Larisa Yarovaya. "Exploring the dynamic relationships between cryptocurrencies and other financial assets." Economics Letters 165 (2018): 28-34.

[4] DeMiguel, Victor, and Francisco J. Nogales. "Portfolio selection with robust estimation." Operations Research 57, no. 3 (2009): 560-577.

[5] Dickey, David A., and Wayne A. Fuller. "Distribution of the estimators for autoregressive time series with a unit root." Journal of the American statistical association 74, no. 366a (1979): 427-431.

[6] Guesmi, Khaled, Samir Saadi, Ilyes Abid, and Zied Ftiti. "Portfolio diversification with virtual currency: Evidence from bitcoin." International Review of Financial Analysis 63 (2019): 431-437.

[7] Granger, Clive W., Esfandiar Maasoumi, and Jeffrey Racine. "A dependence metric for possibly nonlinear processes." Journal of Time Series Analysis 25, no. 5 (2004): 649-669.

[8] Hirschberg, Joseph G., Esfandiar Maasoumi, and Daniel J. Slottje. "Clusters of attributes and well-being in the USA." Journal of Applied Econometrics 16, no. 3 (2001): 445-460.

[9] Hu, Albert S., Christine A. Parlour, and Uday Rajan. "Cryptocurrencies: Stylized facts on a new investible instrument." Financial Management 48, no. 4 (2019): 1049-1068.

[10] Kajtazi, Anton, and Andrea Moro. "The role of bitcoin in well diversified portfolios: A comparative global study." International Review of Financial Analysis 61 (2019): 143-157.

[11] Lahmiri, Salim, and Stelios Bekiros. "Chaos, randomness and multi-fractality in Bitcoin market." Chaos, solitons fractals 106 (2018): 28-34.

[12] Liu, Weiyi. "Portfolio diversification across cryptocurrencies." Finance Research Letters 29 (2019): 200-205.

[13] Lorenzo Álvarez, Luis, and Javier Arroyo Gallardo. "Analysis of the cryptocurrency market applying different prototype-based clustering techniques." Finacial Innovation (2021).

[14] Maasoumi, Esfandiar, and Jeff Racine. "Entropy and predictability of stock market returns." Journal of Econometrics 107, no. 1-2 (2002): 291-312.

[15] Maasoumi, Esfandiar, and Le Wang. "Economic reform, growth and convergence in China." The Econometrics Journal 11, no. 1 (2008): 128-154.

[16] Maasoumi, Esfandiar, and Xi Wu. "Contrasting Cryptocurrencies with Other Assets: Full Distributions and the COVID Impact." Journal of Risk and Financial Management 14, no. 9 (2021): 440.

[17] Maasoumi, Esfandiar, and Xi Wu. 2021. ”Contrasting Cryptocurrencies with Other Assets: Full Distributions and the COVID Impact” Journal of Risk and Financial Management 14, no. 9: 440. https://doi.org/10.3390/jrfm14090440

[18] MacQueen, James. ”Some methods for classification and analysis of multivariate observations.” In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, vol. 1, no. 14, pp. 281-297. 1967.

[19] Mallikarjuna, Mejari, and R. Prabhakara Rao. ”Evaluation of forecasting methods from selected stock market returns.” Financial Innovation 5, no. 1 (2019): 1-16.

[20] Mantegna, Rosario N. ”Hierarchical structure in financial markets.” The European Physical Journal B-Condensed Matter and Complex Systems 11, no. 1 (1999): 193-197.

[21] Nakamoto, Satoshi. ”Bitcoin: A peer-to-peer electronic cash system.” Decentralized Business Review (2008): 21260.

[22] Pele, Daniel Traian, Niels Wesselhöfft, Wolfgang K. Härdle, Michalis Kolossiatis, and Yannis G. Yatracos. ”A statistical classification of cryptocurrencies.” Available at SSRN 3548462 (2020).

[23] Platanakis, Emmanouil, and Andrew Urquhart. ”Should investors include bitcoin in their portfolios? A portfolio theory approach.” The British accounting review 52, no. 4 (2020): 100837.

[24] Platanakis, Emmanouil, Charles Sutcliffe, and Andrew Urquhart. ”Optimal vs naïve diversification in cryptocurrencies.” Economics Letters 171 (2018): 93-96.

[25] Sigaki, Higor YD, Matjaž Perc, and Haroldo V. Ribeiro. ”Clustering patterns in

efficiency and the coming-of-age of the cryptocurrency market." Scientific reports 9, no. 1 (2019): 1-9.

[26] Song, Jung Yoon, Woojin Chang, and Jae Wook Song. "Cluster analysis on the structure of the cryptocurrency market via Bitcoin–Ethereum filtering." Physica A: Statistical Mechanics and its Applications 527 (2019): 121339.

[27] Stosic, Darko, Dusan Stosic, Teresa B. Ludermir, and Tatijana Stosic. "Collective behavior of cryptocurrency price changes." Physica A: Statistical Mechanics and its Applications 507 (2018): 499-509.

[28] Zieba, Damian, Ryszard Kokoszczyński, and Katarzyna Śledziewska. "Shock transmission in the cryptocurrency market. Is Bitcoin the most influential?." International Review of Financial Analysis 64 (2019): 102-125.